CS 524 Homework #2

Jay Kalyanbhai Savani
CWID-20009207

This homework contains both technical and business-related problems, for the total of 100 points.

1. Complete reading Chapter 3 of the textbook **and** the lecture materials. **Please note the errata: The references to [19] on p. 56 of the book should be replaced with references to [20]**!Please also read [20] (available free) at https://www.kernel.org/doc/ols/2007/ols2007v2-pages-87-96.pdf.

Ans: I read the Material

2. **(10 points)** Explain the advantage that paravirtualization provides for handling timers in virtual machines.

Ans: The internal clocks of all contemporary operating systems are kept running via clock interrupts, a function that is especially crucial for real-time media processing. Even an idle virtual computer must handle the clock interruptions for this. Without paravirtualization, the hypervisor would have to plan timer interruptions for idle machines, which is not a scalable method of virtualization. With paravirtualization, the virtual machine code is modified to request a notice at the designated time.

Reference: Textbook : Cloud Computing Business Trends and Technologies Page No: 56

3. **(10 points)** Explain how paravirtualization helps in minimizing access to APIC.

Ans: Only the scheduler and interrupt handlers need to be completely aware of the variations when using a modular architecture; otherwise, an operating system handles several CPUs in the same manner it does one. Without going into too much detail, we should mention that x86-based multi-processor systems employ the Advanced Programmable Interrupt Controller (APIC) to allow symmetric multi-processing by redirecting interrupts (SMP). The transitions into and out of the hypervisor make using APIC in virtual mode costly. A single hypercall may take the place of numerous APIC access requests thanks to paravirtualization, which can see the whole source code.

Reference: Textbook : Cloud Computing Business Trends and Technologies Page No: 56

4. **(5 points)** Find out if *Linux* (like *Unix*) has both the user-mode and system-mode stacks for each process it runs.

Ans: A single Processor may include a large number of identical register sets. A minimum of two register sets must be reserved: one for system mode, also known as supervisory mode or kernel mode, where only the operating system's software is active, and one for user mode, where application applications are executed.

The CPU doesn't go from user mode to system mode automatically. The first thing that happens is that the CPU experiences an interrupt (like timers, keyboards, etc.) When one of these interruptions occurs, the CPU shifts to system mode, pauses the current execution, and then runs the interrupt handler, which activates the system stack pointer. This handler performs its tasks, saves the CPU's status, then restores it before exiting user mode.

Reference: Textbook : Cloud Computing Business Trends and Technologies Page No: 23

5. **(10 points)** Find out what “unscrambled” means in the description of the *Intel LSL* instruction (you can, for example, use the Intel manual referenced in the lecture).

Ans: The term "Unscrambled" in the Intel Load Segment Limit (LSL) command refers to the limit scaled in accordance with the setting of the G flag in the segment descriptor, as per Intel manual, section 5.10.3. When the type and privilege level checks are successful in the destination register and set a ZF flag in the EFLAGS register, the unscrambled limit is loaded. The instruction does not affect the destination register and instead clears the ZF flag if the segment selector is not visible at the current privilege level or is an improper type for the LSL instruction. The processor does a limit check before accessing any segment to make sure the offset is within the segment limit. With the LSL (Load Segment Limit) command, the program may carry out the limit verification. The segment selector for the segment descriptor whose limit is to be verified and a destination register are specified by the LSL instruction. subject to the G flag. Several theories are put forward to explain the constraints. When the G flag is unset, the segment descriptor's 20-bit limit serves as the effective upper limit. When the G flag is set to 4KB page granularity in this case, the limit runs from 0 to 1MB, and the processor scales the value in the limit field by a factor of 212. (4KBytes). The actual limit here is between 4KB and 4GB.

Reference : https://www.cs.cmu.edu/~410/doc/intel-isr.pdf

6. **(20 points)** Read the following two papers:
   - Carl Waldspurger and Rosenblum, M. (2012) *I/O Virtualization.* Communications of the ACM, vol. 55, No 1. January 2012. Pages 66-72; and
   - Muli Ben-Yehuda; Xenidis, J.; Ostrowski, M.; Rister, K.; Bruemmer, A.; Van Doorn, L. (2007). *The Price of Safety: Evaluating IOMMU Performance.* Proceedings of the Linux Symposium on June 27th–30th, 2007. Ottawa, Ontario. Pages 225-230.

   1) Explain the advantages and disadvantages of using I/O MMU by citing the appropriate text from the paper;

Ans: Sol 1) **Benefits:**

- The IOMMU translates contiguous virtual addresses to the underlying fragmented physical addresses, enabling large sections of memory to be allocated without the need that they be contiguous in physical memory. As a result, it's occasionally possible to avoid using vectored I/O (scatter- gather lists).
- Devices that lack the memory address lengths necessary to target the complete physical memory may nonetheless access it via the IOMMU, saving on the costs involved with transferring buffers to and from the peripheral's accessible memory area.
- For instance, the Physical Address Extension (PAE) feature in an x86 CPU allows x86 computers to access memory that is larger than 4 gigabytes. Yet, a typical 32-bit PCI device cannot directly access RAM that is over the 4 GiB border since it cannot address it. The operating system would have to incorporate time-consuming bounce buffers in the absence of an IOMMU.
- Since a device cannot read from or write to memory that has not been specifically allocated (mapped) for it, memory is safeguarded against malicious devices that try DMA assaults and malfunctioning devices that attempt erroneous memory transfers.
- The memory protection is provided by the MMU and IOMMU being under the exclusive control of the OS executing on the CPU (see figure). Physically, the devices cannot tamper with or damage specified memory management tables. Guest operating systems in virtualization might make use of hardware that isn't designed for virtualization.
- Use of high-performance gear, including graphics cards
- DMA devices fail in virtual environments because the virtual machine software remaps all memory locations, making it impossible to access memory directly. This remapping is handled by the IOMMU, enabling the usage of native device drivers in a guest operating system.
- Like normal memory address re-mapping, IOMMU also does hardware interrupt re-mapping in certain architectures. An IOMMU may provide peripheral memory paging. Memory management services may be detected and requested by a peripheral utilizing the PCI-SIG PCIe Address Translation Services (ATS) Page Request Interface (PRI) extension.

Drawbacks:

- Possible performance reduction due to translation and administration overhead. Physical memory use for the new I/O page translation tables. If the tables can be shared with the CPU, this may be lessened.

Reference - https://en.wikipedia.org/wiki/Input%E2%80%93output_memory_management_unit

6th Question Cont.

Solution 2) Virtualization services that satisfy some or all of the anticipated qualities included in carrier grade solutions are referred to as carrier grade virtualization (CGV). Now let's look at these characteristics and the demands they place on carrier grade virtualization. While not a full description, the list is detailed enough to explain carrier grade virtualization.

- Availability: Virtualization may raise carrier-grade systems' overall availability, preserving or even improving the 5 to 7 nines of service that are now anticipated and given (through lengthened MTTF or shortened MTTR).
- High Performance Scaling: Carrier Grade Virtualization must support scalability by enabling the capabilities of new hardware, notably multicore CPUs, and by operating at high efficiency with little overhead.
- Minimal error recovery domains: Virtualization separates virtual machines from their guest operating systems; if one VM or guest OS fails, it has no effect on other collocated VMs. By designating virtual machines (VMs) to particular tasks like granting device access, carrier grade virtualization takes use of this "natural" property and enhances it.
- Carrier Grade Virtualization must be able to handle real-time workloads and systems, hence it must have real-time and deterministic features. This is because software components used in carrier grade systems do have real-time requirements.
- Upgrade capabilities: OSes, middleware, and application components are now instantiated on hardware platforms differently as a result of virtualization. Software updates for such components must be made easier by carrier grade virtualization, which must also facilitate selfupgrade. Figure 2 shows a simplified architectural schematic for the SCOPE Alliance Carrier Grade Platform. VirtualLogix Real-Time VirtualizationTM o Configurable security: Virtualization separates virtual machines and their workloads from one another. 7 VL/TR-08-132.1. Carrier Grade Virtualization dictates that security is not handled in a "one size fits all" manner. To govern VM access to real resources and establish and implement suitable security rules across virtual machines, CGV must have these features (CPU, memory, devices...).
- Effective and uniform management interfaces: The introduction of virtual machines through virtualization creates new objects that need to be handled. The management interfaces for carrier grade virtualization must be effective and consistent.

Reference : www.linuxpundit.com/documents/CGV_WP_Final_FN.pdf

7. **(5 points)** Find out what hypervisors *Amazon* is using in EC2, and describe their major characteristics.

Ans: Xen Automotive: Its automobile hypervisor avoids the expense of creating, locating, installing, and testing several separate processors within the car. It supports all infotainment operations in the vehicle inside a single compute engine. Real-time streaming of numerous films, virtualization of GPUs, driver support for specialized hardware, and new guest operating systems are just a few of the many problems it poses. It would be challenging and probably fall short of the aim to simply repurpose a hypervisor connected to a conventional operating system, but similar adjustments are easily accomplished on a bare-metal hypervisor.

• Featured are:

- Live VM Migration: It facilitates moving virtual machines live across hosts, which helps to balance workloads and prevent downtime.
- Live Storage Migration: Move virtual machines that are currently in use, together with the virtual disk images that go with them, inside and across resource pools by using both local and shared storage.
- HostFailureProtection: Provide high availability by automating virtual machine restarts in the event of a server, hypervisor, or VM failure. Network interfaces are joined via link aggregation to boost network performance and provide redundancy.
- HostPowerProtection: Use built-in hardware characteristics to reduce the amount of energy used by datacenters by dynamically consolidating virtual machines onto fewer systems and shutting down idle servers when service demand changes.

- Memory Overcommit: By distributing unused server memory across the virtual machines running on the host server, you may save expenses while enhancing application speed and security.
- Site Recovery: Offers services and planning for site-to-site catastrophe recovery in virtual settings. Site recovery is simple to set up, quick to recover from, and can be tested regularly to make sure disaster recovery strategies are still effective.
- Density: Workloads have plenty of space thanks to a razor-thin hypervisor. Yet because to cutting-edge innovations like Unkennels, a huge number of visitors may be handled on a single server.
- Scalability: A bare-metal hypervisor's capacity to scale is independent of the host operating system. It can be modified to make the most use of the resources available in a future with thousands of very modest workloads.
- Security: The absence of a host operating system reduces the attack surface for malevolent hackers. Security is a major issue that is only becoming worse in the realm of the cloud. It is incredibly beneficial to not have a host operating system to exploit.
- Scheduling: The hypervisor is free to utilize any scheduler the workload warrants. The host operating system's scheduler, which may not be the best option for the workloads tomorrow, is not forced upon you.
- Paravirtualization: Due to paravirtualization's streamlined interface, specialized workloads don't need to create intricate drivers for simulated hardware.

Reference: https://phoenixnap.com/blog/what-is-bare-metal-hypervisor

8. **(5 points)** Find out the URL to the source code of the Nitro hypervisor

Ans: The user guide for the Nitro Enclaves CLI can be found at https://docs.aws.amazon.com/enclaves/latest/user/nitro-enclave-cli.html.

Ensure that your EC2 instance was created with enclave support enabled and that your system (*and container if applicable*) has read/write access to /dev/nitro_enclaves.
Ensure that your Linux system (*and container if applicable*) has Linux hugepages available.

The AWS Nitro Enclaves CLI package is currently available for:

- Amazon Linux 2 - https://docs.aws.amazon.com/enclaves/latest/user/nitro-enclave-cli-install.html
- openSUSE and SUSE Linux Enterprise Server - https://build.opensuse.org/package/show/Cloud:Tools/aws-nitro-enclaves-cli
- Windows - https://docs.aws.amazon.com/enclaves/latest/user/nitro-enclave-cli-install-win.html

Enclave disk size

The enclaves do not have access to a physical disk, just a RAM filesystem. One can configure the disk space by changing memory size or by using kernel command line arguments.

The init.c file keeps the default configuration for each volume. The below example shows the default options for /tmp.

```
{ OpMount, .mount = { "tmpfs", "/tmp", "tmpfs", MS_NODEV | MS_NOSUID | MS_NOEXEC } },
```

To modify the memory allocated to this volume, another parameter is needed

```
{ OpMount, .mount = { "tmpfs", "/tmp", "tmpfs", MS_NODEV | MS_NOSUID | MS_NOEXEC,
"size=100%" } },
```

Note that the parameter size specifies only the maximum allocated size. After modifying the configuration, the file needs to be recompiled using make init and moved to /usr/share/nitro_enclaves/blobs/init.

Reference - https://docs.aws.amazon.com/enclaves/latest/user/nitro-enclave-cli.html.

9. **(10 points)** Examine the *Amazon* EC2 VM offer capabilities and particularly the Amazon Machine Image (AMI) ((https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AMIs.html) and answer the following questions:

    a. How (i.e., in what units) does EC2 measure the CPU power of a virtual machine and how is the unit in question translated into the power of the physical processors)?
    b. Which operating systems are available on the above systems?
    c. What is an AMI and what is its relationship to an *instance*?
    d. What are the components of an AMI?

Ans: Amazon EC2 platform is available for free to try. There are no hidden or minimum charges and the user must pay according to the need and the usage of the services. And to calculate the estimated monthly bill, a monthly calculator available. There are four ways to pay for Amazon EC2 instances: On-Demand, Reserved Instances, Spot Instances and Dedicated Hosts.

- On-DemandInstanceisusefulforunpredictedorshort-termworkloadforthisuser must pay for compute capacity by the hour with no long-term commitments or upfront payments.
- ReservedInstancesareusefulforpredictableworkloadtoreservethecapacityand available with significant discounts up to 75% off as compared to On-Demand Instance. All the standard reserved instances are available always (i.e., 24x7) and allow launching the reserved instances at the time of need.
- Spot Instances are helpful for urgent computing needs for large amounts of additional capacity and allow to bid on spare computing capacity for up to 90% off the On-Demand Instances price.
- Dedicated Hosts are helpful to meet compliance requirements to reduce costs by allowing using existing server-bound software licenses (subject to user's license terms). Dedicated Hosts either can be purchased OnDemand on an hourly basis or can be purchased as a reservation for up to 70% off the On-Demand price.

10. **(10 points)** Find out about the pricing of the EC2 platforms and provide a few examples.

Ans: • The Amazon EC2 platform is available for free to try. There are no hidden or minimum charges and the user have to pay according to the need and the usage of the services. And to calculate the estimated monthly bill, a monthly calculator available.There are four ways to pay for Amazon EC2 instances: On-Demand, Reserved Instances, Spot Instances and Dedicated Hosts.

- On-Demand Instance is useful for unpredicted or short-term workload for this user has to pay for compute capacity by the hour with no long-term commitments or upfront payments.
- Reserved Instances are useful for predictable workload to reserve the capacity and available with significant discounts up to 75% off as compared to On-Demand Instance. All the standard reserved instances are available always (i.e. 24x7) and allow launching the reserved instances at the time of need.
- Spot Instances are helpful for urgent computing needs for large amounts of additional capacity and allow to bid on spare computing capacity for up to 90% off the On-Demand Instances price.
- Dedicated Hosts are helpful to meet compliance requirements to reduce costs by allowing using existing server-bound software licenses (subject to user's license terms). Dedicated Hosts either can be purchased On-Demand on an hourly basis or can be purchased as a reservation for up to 70% off the On-Demand price.

References: https://aws.amazon.com/ec2/pricing/

Instance Name	RI upfront fee	RI Monthly fee	Hourly Rate	Savings	On-demand Rate
A1.medium	$0	$11.75	$0.016	37%	0.255
T4g.small	$0	$07.67	$0.011	37%	0.0168
T3.nano	$0	$02.41	$0.0103	37%	0.0052
T3.2xlarge	$0	$152.30	$0.209	37%	0.3328

11. **(15 points)** From the above exercise, you will learn that it is possible to create a free machine instance. Please, do the following:

    a. Find out and document the essence of the respective *Service Level Agreement (SLA)* on*;* in particular write down what one needs to do in order to maintain this service **free;**

    b. Describe the process (i.e., what exactly one needs to do) to create a free machine instance that could be used as a server. (**Do not,** however, create anything yet!)

    c. Can you create a machine instance equivalent to your own PC and then transfer your own PC image there? If so, how would you achieve that?

Ans : a). service Level Agreement (SLA) is a contract between a cloud provider (either internal or external) and the service user that outlines responsibilities, quality, and scope on both sides. The most common component of SLA is that the services should be provided to the customer as agreed upon in the contract. To maintain free services of Amazon EC2, one needs to sign up under the Free Tier, to get hands on experience for 12-month period. Then the one need to create an account and use the services provided under certain usage limits.

The need to follow the steps:

- Sign up for an AWS account.
- Must provide credit card information and billing address. Until the free usage exceeds the limits, you would not be charged for the services.
- Get started with AWS Cloud services by choosing any of the products listed under the Free Tier service.

References: https://www.techtarget.com/searchcloudcomputing/

b). The process what exactly one need to do to create a free machine instance, that could be used as a server are followed: i. First, must create an instance of Amazon EC2 which can be used as a server for hosting an application on the cloud. ii. Then must create a server for the database which would be a database instance. iii. After performing above steps, a web app can be deployed on the server. iv. After that, load balancing and scaling needs to be done so that the traffic is distributed across the number of servers or application servers. v. In the last, user can associate or use a name with your web application.

References: http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AMIs.html

c). Yes, we can create a machine instance equivalent to my own PC and then transfer our own PC image there. All of this can be done by creating an EC2 instance on the Amazon Cloud and host it as a server. After that, we need to connect our own PC to that server and then transfer the image.

References: https://aws.amazon.com/premiumsupport/knowledge-center/