

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-8796.pdf>.

Ans: I have completed reading chapter 3 of the textbook and the lecture materials.

(Reference: Cloud Computing: Business Trends and Technologies)

2. Explain the advantage that paravirtualization provides for handling timers in virtual machines.

Ans: In Cloud Computing, Para-Virtualization is a virtualization technique that presents a software interface to the virtual machines which is similar, yet not identical to the underlying hardware software interface. All modern operating systems, even idle virtual machines rely on clock interrupts to maintain their internal timers, a feature that is particularly essential for real time media processing. With para-virtualization, a small modification is to change the idle code to request the VMM to notify itself in a specified time period. Again, time is calculated and restored in the guest. The hypervisor without the paravirtualization would need to schedule timer interrupts back-to-back for idle machines when guest operating system is scheduled back to run, which would not be considered scalable way of virtualization.

(Reference: Cloud Computing: Business Trends and Technologies)

3. Explain how paravirtualization helps in minimizing access to APIC.

Ans: An operating system deals with multiple CPUs same way it deals with modular design, it is just scheduler and interrupt handlers that need to be aware of the differences. The code sending IPI on the x86-64 native systems using the flat mode needs to access the APIC registers a couple of times. Each access to the APIC registers needs to be intercepted for virtualization, causing overhead. With the help of para-virtualization multiple requests which are implicit with a single explicit hypercall, achieving faster, simpler and more efficient implementations. Paravirtualization determines the full view of the code which indeed access the multiple APIC.

(Reference: Cloud Computing: Business Trends and Technologies)

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

Ans: Yes, in Linux like Unix, which has both user mode and kernel mode (system mode) stacks for each process it runs.

The CPU may have more than one set of identical registers. The least the operating system will work is to set one set of register reserved for user mode where application program executes and other for the system mode in which only the operating system software executes.

The switch between the modes will not be done automatically by the CPU. The interrupt handler switches between the user mode and system mode when interrupts occur in the CPU.

(Reference: Cloud Computing: Business Trends and Technologies)

5. 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: With Intel Manual section 3.2 Instruction Set Reference, A-L, Pg. 646, “Unscrambled” in Intel load Segment Limit (LSL) instruction means the ideal limit scaled according to the setting of the G flag in the segment descriptor. When the privilege level and type checks pass into destination register and set a ZF flag in the EFLAGS register then the unscrambled limit is loaded. If the segment selector is not visible at the current privilege level or is an invalid type for the LSL instruction, the instruction does not modify the destination register and clears the ZF flag. Its due to the fact that the limit field is spread across several bits within GDT entry.

The software can perform the limit checking using the LSL (LOAD SEGMENT LIMIT) instruction. The LSL instruction specifies the segment selector for the segment descriptor whose limit is to be checked and destination register. Based on the G flag, the limits are interpreted. When the G flag is clear, the effective limit is the value of the 20-bit limit in the segment descriptor. The limit ranges from 0 to 1 MB, when the G flag is set 4KB page granularity, the processor scales the value in the limit field by a factor of 212(4000 bytes). It ranges from 4000bytes to 4GB.

(Reference: <https://www.intel.com/content/www/us/en/architecture-and-technology/64-ia-32-architectures-software-developer-instruction-set-reference-manual-325383.html>)

6. 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.

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

6.2) Research the Web to find what is meant by “carrier-grade hypervisors”. What products are available?

Ans 6.1:

Advantages of I/O MMU

- I/O MMU translates the I/O virtual memory address to corresponding physical memory, making direct access by devices safe and efficient and allows the driver in the VM to program device DMA using its virtual memory address.
- Devices that do not support memory address long enough to address the entire physical memory can still address the entire memory through the I/O MMU, avoiding overheads associated with copying buffers to and from the peripheral’s addressable memory space.
- The ability to multiplex logical I/O devices, allowing multiple logical devices to be implemented by a smaller number of physical devices.
- VM features such as the ability to suspend and resume a VM and the ability to move a running VM between physical machines known as live migration.
- This virtualization layer may also change mappings to physical devices even when the VM does not move.

f. One useful capability enabled by I/O virtualization is device aggregation where multiple physical devices can be combined into a single more capable logical device that is exported to the VM.

g. By interposing and transforming virtual I/O requests transparently enhancing unmodified software with new capabilities, features can be added to existing systems.

Dis-Advantages of I/O MMU

a. First and foremost, it is only applicable to a hypervisor scenario. In a bare-metal scenario, getting rid of map and uncap isn't practical because it renders the I/O MMU useless—if one maps all physical memory, why use an IOMMU at all?

b. Second, even in a hypervisor scenario, pre-allocation is only viable if the set of machine frames owned by the guest is “mostly constant” through the guest's lifetime. If the guest wishes to use page flipping or ballooning, or any other operation which modifies the guest's pseudo-physical to machine mapping, the I/O MMU mapping needs to be updated as well so that the I/O to machine mapping will again correspond exactly to the pseudo-physical to machine mapping.

c. Another downside of this optimization is that it protects other guests and the hypervisor from the guest, but provides no protection inside the guest itself.

(Reference: <https://cacm.acm.org/magazines/2012/1/144808-i-o-virtualization/fulltext> , I/O Virtualization by Carl Waldspurger, Mendel Roseblum, Wikipedia Input Output Memory management unit)

Ans 6.2:

Carrier Grade can be defined as virtualization services that fulfil some or all expected properties existing in edge and core network elements such as IP Multimedia Systems nodes.

Features include Real time Behaviour, configurable security, Availability, High performance scaling, Upgrade capabilities, fault tolerance and easy analysis.

Products available are: BareMetal Xen Hypervisors, Oracle Solaris, NEC CGHV.

7. Find out what hypervisors Amazon is using in EC2 and describe their major characteristics.

Ans: Amazon EC2 uses bare metal hypervisors in Xen.

Major characteristics are:

a. Live Storage Migration: It supports virtual machine running and their associated virtual disk image within and across resource tools and leveraging local and shared storage.

b. Live VM Migration: Supports migration from one host to another allows workload balancing and the avoidance of downtime.

c. Host power protection: Take advantage of embedded hardware features to lower data centre electricity consumption by dynamically consolidating by VMs on fewer systems and then powering off underutilized servers as demand for services fluctuates.

d. Host failure protection: Deliver high availability by automatically restarting virtual machines if a failure occurs at VM hypervisor level.

e. Memory Overcommit: Helps improve application performance and reduce costs and protection by sharing unused server memory between VMs on the host server.

(References: https://en.wikipedia.org/wiki/Amazon_Elastic_Compute_Cloud)

8. 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. What kinds of machine instances are there as characterized by the power of their respective CPUs, platform (i.e., 32-bit or 64-bit), memory, storage, etc.? Please list all the instances in the nomenclature along with their respective characteristics.

c. Which operating systems are available on the above systems?

d. What is an AMI and what is its relationship to an instance?

e. What are the components of an AMI?

Ans:

8.a. The amount of CPU that is allocated to a particular instance is expressed in terms of these EC2 Compute Units. We use several benchmarks and tests to manage the consistency and predictability of the performance from an EC2 Compute Unit. One EC2 Compute Unit provides the equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor. This is also the equivalent to an early-2006 1.7 GHz Xeon processor referenced in our original documentation.

(Reference:

https://aws.amazon.com/ec2/faqs/#What_is_an_EC2_Compute_Unit_and_why_did_you_introduce_it)

8.b. Amazon EC2 provides a wide selection of different instance types and provides the flexibility to choose the combination of instance to meet the computing need most appropriately and these sets of instance combinations can be changed later depending upon change in business need.

The various types of instances are:

General Purpose:

This instance family includes A1, T2, M5, and M5a,T3 which is often the first choice because of variety of CPU size range.

1. Amazon EC2 A1 instances deliver significant cost savings and are ideally suited for scale-out and Arm-based workloads that are supported by the extensive Arm ecosystem. A1 instances are the first EC2 instances powered by AWS Graviton Processors that feature 64-bit Arm Neoverse cores and custom silicon designed by AWS.

Features:

Custom built AWS Graviton Processor with 64-bit Arm Neoverse cores

Support for Enhanced Networking with Up to 10 Gbps of Network bandwidth EBS-optimized by default

2. T2 instances are Burstable Performance Instances that provides a baseline level of CPU performance with the ability to burst above the baseline, which is governed by CPU Credits.

These instances are good choices for workloads that don't use the full CPU often or consistently, but occasionally need to burst (examples web servers, development environments and databases).

Features:

- High frequency Intel Xeon processors
- Burstable CPU, governed by CPU credits, and consistent baseline performance
- Lowest-cost general purpose instance type, and Free Tier eligible (t2.micro only)
- Balance of computing, memory and network resources

3. M4 instances are the latest generation of General-Purpose Instances. This provides a balance of computing, memory and network resources, and it is a good choice for many applications.

Features:

- 2.3 GHz Intel Xeon E5-2686 v4 (Broadwell) processors or 2.4 GHz Intel Xeon E5-2676 (Haswell) processors
- EBS-optimized by default at no additional cost
- Support for Enhanced Networking
- Balance of computing, memory, and network resources

Compute Optimized: C4, C5, C5n are the compute optimized instances.

1. C5 instances are optimized for compute-intensive workloads and deliver cost-effective high performance at a low price per compute ratio.

Features:

- 3.0 GHz Intel Xeon Platinum processors with Intel Advanced Vector Extension 512 (AVX-512) instruction set.
- Run each core at up to 3.5 GHz using Intel Turbo Boost Technology
- Larger instance size, c5.18xlarge, offering 72 vCPUs and 144 GiB of memory
- Requires HVM AMIs that include drivers for ENA and NVMe

2. C4 instances are optimized for compute-intensive workloads and deliver very cost-effective high performance at a low price per compute ratio.

Features:

High frequency Intel Xeon E5-2666 v3 (Haswell) processors optimized specifically for EC2
Default EBS-optimized for increased storage performance at no additional cost

Higher networking performance with Enhanced Networking supporting Intel 82599 VF.

Memory Optimized: R5, R5a, X1, X1e, R4 are the instances

1. R4 instances are optimized for memory intensive applications and offer lower price per GiB of RAM than R3.

Features:

- 2.3 GHz Intel Xeon E5-2686 v4 (Broadwell) processors
- DDR4 Memory

2. R5 instances deliver 5% additional memory per vCPU than R4 and the largest size provides 768 GiB of

memory. In addition, R5 instances deliver a 10% price per GiB improvement and a ~20% increased CPU performance over R4.

Features:

- Up to 3.1 GHz Intel Xeon® Platinum 8175 processors with new Intel Advanced Vector Extension (AVX-512) instruction set
- Up to 768 GiB of memory per instance
- Powered by the AWS Nitro System, a combination of dedicated hardware and lightweight hypervisor

3. X1 instances are optimized for large-scale, enterprise-class and in-memory applications, and offer one of the lowest price per GiB of RAM among Amazon EC2 instance types.

Features:

- High frequency Intel Xeon E7-8880 v3 (Haswell) processors
- One of the lowest prices per GiB of RAM
- Up to 1,952 GiB of DRAM-based instance memory

Accelerated Computing: P2, P3 G3, F1 are the instances available.

1. P2 instances are intended for general-purpose GPU compute applications.

Features:

- High frequency Intel Xeon E5-2686 v4 (Broadwell) processors
- High-performance NVIDIA K80 GPUs, each with 2,496 parallel processing cores and 12GiB of GPU memory
- Supports GPUDirect™ for peer-to-peer GPU communications

2. P3 instances are the latest generation of general-purpose GPU instances.

Features:

- Up to 8 NVIDIA Tesla V100 GPUs, each pairing 5,120 CUDA Cores and 640 Tensor Cores
- High frequency Intel Xeon E5-2686 v4 (Broadwell) processors for p3.2xlarge, p3.8xlarge, and p3.16xlarge.
- High frequency 2.5 GHz (base) Intel Xeon P-8175M processors for p3dn.24xlarge.
- Supports NVLink for peer-to-peer GPU communication

3. G3 instances are optimized for graphics-intensive applications.

Features:

- High frequency Intel Xeon E5-2686 v4 (Broadwell) processors
- NVIDIA Tesla M60 GPUs, each with 2048 parallel processing cores and 8 GiB of video memory
- Enables NVIDIA GRID Virtual Workstation features, including support for 4 monitors with resolutions

up to 4096x2160. Each GPU included in your instance is licensed for one "Concurrent Connected User"

- Enables NVIDIA GRID Virtual Application capabilities for application virtualization software like Citrix

XenApp Essentials and VMware Horizon, supporting up to 25 concurrent users per GPU

Storage Optimized: This instance family includes the H1, I3 and D2 instance types, and provides you with direct attached storage options optimized for applications with specific disk I/O and storage capacity requirements.

- I3 also offers Bare Metal instances (i3.metal), powered by the Nitro System, for non-virtualized workloads, workloads that benefit from access to physical resources, or workloads that may have license restrictions.

Features:

- High Frequency Intel Xeon E5-2686 v4 (Broadwell) Processors with base frequency of 2.3 GHz
 - Up to 25 Gbps of network bandwidth using Elastic Network Adapter (ENA)-based Enhanced Networking
 - High Random I/O performance and High Sequential Read throughput
- D2 – Dense Storage instances feature up to 48TB of HDD-based local storage, deliver high disk throughput, and offer the lowest price per disk throughput performance on Amazon EC2.

Features:

- High-frequency Intel Xeon E5-2676v3 (Haswell) processors
- HDD storage
- Consistent high performance at launch time
- High disk throughput
- Support for Amazon EC2 Enhanced Networking

3.H1 instances feature up to 16 TB of HDD-based local storage, deliver high disk throughput, and a balance of compute and memory.

Features:

- Powered by 2.3 GHz Intel® Xeon® E5 2686 v4 processors (codenamed Broadwell)
- Up to 16TB of HDD storage
- High disk throughput
- ENA enabled Enhanced Networking up to 25 Gbps

(References: <https://aws.amazon.com/blogs/aws/choosing-the-right-ec2-instance-type-for-your-application/>, <https://aws.amazon.com/ec2/instance-types/>)

8.c. Amazon EC2 supports almost all operating systems.

Some of them are:

Amazon Linux, CentOS, CoreOS, Debian, Fedora, FreeBSD, Genymotion, Oracle Linux, RancherOS, Red Hat Enterprise Linux (RHEL), SUSE, SUSE Linux Enterprise Server, TurnKey Core, Windows Server, and Ubuntu Server 18.08 LTS

(Reference <https://docs.aws.amazon.com/opsworks/latest/userguide/workinginstances-os.html>)

8.d. An instance is a virtual machine with particular specifications and OS that you choose while creating them. An AMI (Amazon Machine Image) is a complete backup of an instance. When you make AMI of an instance, two things happen

1. AMI Creation that has all its launch configurations and
2. Snapshot attached to this AMI which has disk backup of the instance.

8.e. Each Amazon Machine Image includes the following:

- A template for the instance's root device volume
- Set up permissions that control which Amazon Web Services (AWS) accounts can use the machine images to set up an instance
- The block devices that specify the root device volume to attach to the E2C instance once it's launched.

A read-only filesystem image - the main component of a machine image - is normally compressed, encrypted and uploaded into [Amazon S3](#) for storage. The file used to store information about the machine image is known as a manifest file.

Information stored in this file includes name, architecture, decryption key and default kernel id. Note that an AMI does not include an image but rather a pointer to the default kernel id.

(Reference: <https://www.linkeit.com/blog/what-are-amazon-machine-images-ami>)

9. 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.

There are four ways to pay for Amazon EC2 instances:

On-Demand, Reserved Instances, Spot Instances and Dedicated Hosts.

1. With On-Demand instances, you pay for compute capacity by per hour or per second depending on which instances you run. No longer-term commitments or upfront payments are needed. You can increase or decrease your compute capacity depending on the demands of your application and only pay the specified per hourly rates for the instance you use.

On-Demand instances are recommended for:

- Users that prefer the low cost and flexibility of Amazon EC2 without any up-front payment or long-term commitment
- Applications with short-term, spiky, or unpredictable workloads that cannot be interrupted
- Applications being developed or tested on Amazon EC2 for the first time.

2. Amazon EC2 Spot instances allow you to request spare Amazon EC2 computing capacity for up to 90% off

the On-Demand price.

Spot instances are recommended for:

- Applications that have flexible start and end times
- Applications that are only feasible at very low compute prices
- Users with urgent computing needs for large amounts of additional capacity

3. Reserved Instances provide you with a significant discount (up to 75%) compared to On-Demand instance pricing. In addition, when Reserved Instances are assigned to a specific Availability Zone, they

provide a capacity reservation, giving you additional confidence in your ability to launch instances when

you need them. For applications that have steady state or predictable usage, Reserved Instances can provide significant savings compared to using On-Demand instances. Reserved Instances are recommended for:

- Applications with steady state usage
- Applications that may require reserved capacity
- Customers that can commit to using EC2 over a 1 or 3 year term to reduce their total computing costs

4. A Dedicated Host is a physical EC2 server dedicated for your use. Dedicated Hosts can help you reduce costs by allowing you to use your existing server-bound software licenses, including Windows Server, SQL Server, and SUSE Linux Enterprise Server (subject to your license terms), and can also help you meet compliance requirements.

- Can be purchased On-Demand (hourly).
- Can be purchased as a Reservation for up to 70% off the On-Demand price.

Examples:

On-Demand pricing: For On-demand instances, you need to pay for compute capacity used per hour or

minute with no long-term commitments.

The pricing below includes the cost to run public and private AMIs on the specified operating system.

On-Demand Pricing

On-Demand instances let you pay for compute capacity by the hour or second (minimum of 60 seconds) with no long-term commitments. This frees you from the costs and complexities of planning, purchasing, and maintaining hardware and transforms what are commonly large fixed costs into much smaller variable costs.

The pricing below includes the cost to run private and public AMIs on the specified operating system ("Windows Usage" prices apply to Windows Server 2003 R2, 2008, 2008 R2, 2012, 2012 R2, 2016, and 2019). Amazon also provides you with additional instances for Amazon EC2 running Microsoft Windows with SQL Server, Amazon EC2 running SUSE Linux Enterprise Server, Amazon EC2 running Red Hat Enterprise Linux and Amazon EC2 running IBM that are priced differently.

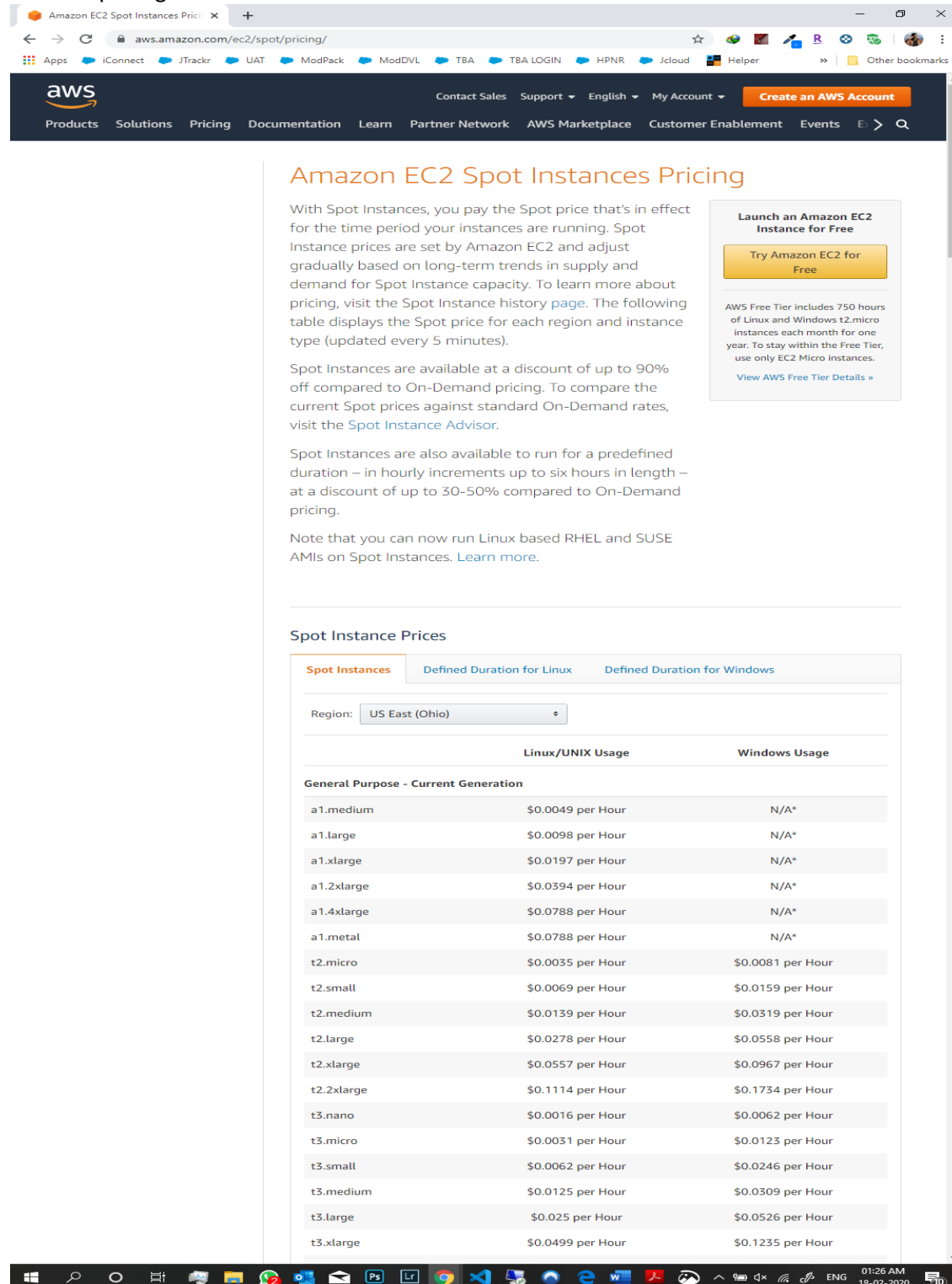
Region: **US East (Ohio)**

	vCPU	ECU	Memory (GiB)	Instance Storage (GB)	Linux/UNIX Usage
General Purpose - Current Generation					
a1.medium	1	N/A	2 GiB	EBS Only	\$0.0255 per Hour
a1.large	2	N/A	4 GiB	EBS Only	\$0.051 per Hour
a1.xlarge	4	N/A	8 GiB	EBS Only	\$0.102 per Hour
a1.2xlarge	8	N/A	16 GiB	EBS Only	\$0.204 per Hour
a1.4xlarge	16	N/A	32 GiB	EBS Only	\$0.408 per Hour
a1.metal	16	N/A	32 GiB	EBS Only	\$0.408 per Hour
t3.nano	2	Variable	0.5 GiB	EBS Only	\$0.0052 per Hour
t3.micro	2	Variable	1 GiB	EBS Only	\$0.0104 per Hour
t3.small	2	Variable	2 GiB	EBS Only	\$0.0208 per Hour
t3.medium	2	Variable	4 GiB	EBS Only	\$0.0416 per Hour
t3.large	2	Variable	8 GiB	EBS Only	\$0.0832 per Hour
t3.xlarge	4	Variable	16 GiB	EBS Only	\$0.1664 per Hour
t3.2xlarge	8	Variable	32 GiB	EBS Only	\$0.3328 per Hour
t3a.nano	2	Variable	0.5 GiB	EBS Only	\$0.0047 per Hour
t3a.micro	2	Variable	1 GiB	EBS Only	\$0.0094 per Hour
t3a.small	2	Variable	2 GiB	EBS Only	\$0.0188 per Hour
t3a.medium	2	Variable	4 GiB	EBS Only	\$0.0376 per Hour
t3a.large	2	Variable	8 GiB	EBS Only	\$0.0752 per Hour
t3a.xlarge	4	Variable	16 GiB	EBS Only	\$0.1504 per Hour
t3a.2xlarge	8	Variable	32 GiB	EBS Only	\$0.3008 per Hour
t2.nano	1	Variable	0.5 GiB	EBS Only	\$0.0058 per Hour
t2.micro	1	Variable	1 GiB	EBS Only	\$0.0116 per Hour
t2.small	1	Variable	2 GiB	EBS Only	\$0.023 per Hour
t2.medium	2	Variable	4 GiB	EBS Only	\$0.0464 per Hour
t2.large	2	Variable	8 GiB	EBS Only	\$0.0928 per Hour
t2.xlarge	4	Variable	16 GiB	EBS Only	\$0.1856 per Hour
t2.2xlarge	8	Variable	32 GiB	EBS Only	\$0.3712 per Hour

For Linux operating system, Amazon EC2 on-demand instance in Ohio region is listed above for General purpose- Current generation.

Amazon EC2 Spot instances: In spot instances, you pay for the spot price for the time your instances are

running. Discounts on spot instances are available up to 90% when the prices are compared with On-demand pricing.



Amazon EC2 Spot Instances Pricing

With Spot Instances, you pay the Spot price that's in effect for the time period your instances are running. Spot Instance prices are set by Amazon EC2 and adjust gradually based on long-term trends in supply and demand for Spot Instance capacity. To learn more about pricing, visit the [Spot Instance history page](#). The following table displays the Spot price for each region and instance type (updated every 5 minutes).

Spot Instances are available at a discount of up to 90% off compared to On-Demand pricing. To compare the current Spot prices against standard On-Demand rates, visit the [Spot Instance Advisor](#).

Spot Instances are also available to run for a predefined duration – in hourly increments up to six hours in length – at a discount of up to 30-50% compared to On-Demand pricing.

Note that you can now run Linux based RHEL and SUSE AMIs on Spot Instances. [Learn more.](#)

Launch an Amazon EC2 Instance for Free
[Try Amazon EC2 for Free](#)
AWS Free Tier includes 750 hours of Linux and Windows t2.micro instances each month for one year. To stay within the Free Tier, use only EC2 Micro instances.
[View AWS Free Tier Details »](#)

Spot Instance Prices

Spot Instances Defined Duration for Linux Defined Duration for Windows

Region:

	Linux/UNIX Usage	Windows Usage
General Purpose - Current Generation		
a1.medium	\$0.0049 per Hour	N/A*
a1.large	\$0.0098 per Hour	N/A*
a1.xlarge	\$0.0197 per Hour	N/A*
a1.2xlarge	\$0.0394 per Hour	N/A*
a1.4xlarge	\$0.0788 per Hour	N/A*
a1.metal	\$0.0788 per Hour	N/A*
t2.micro	\$0.0035 per Hour	\$0.0081 per Hour
t2.small	\$0.0069 per Hour	\$0.0159 per Hour
t2.medium	\$0.0139 per Hour	\$0.0319 per Hour
t2.large	\$0.0278 per Hour	\$0.0558 per Hour
t2.xlarge	\$0.0557 per Hour	\$0.0967 per Hour
t2.2xlarge	\$0.1114 per Hour	\$0.1734 per Hour
t3.nano	\$0.0016 per Hour	\$0.0062 per Hour
t3.micro	\$0.0031 per Hour	\$0.0123 per Hour
t3.small	\$0.0062 per Hour	\$0.0246 per Hour
t3.medium	\$0.0125 per Hour	\$0.0309 per Hour
t3.large	\$0.025 per Hour	\$0.0526 per Hour
t3.xlarge	\$0.0499 per Hour	\$0.1235 per Hour

Amazon EC2 Reserved Instances: When compared to on-demand instance, reserved instances provide you with a significant discount up to 75%. When you purchase a standard or convertible reserved instance, you have three options to pay: All upfront, Partial upfront, No upfront options.

Amazon EC2 Standard Reserved Instance Payment Options

EC2 Instance Pricing – Amazon V

aws.amazon.com/ec2/pricing/reserved-instances/pricing/

Apps iConnect JTrackr UAT ModPack ModDVL TBA TBA LOGIN HPNR Jcloud Helper Other bookmarks

aws

Contact Sales Support English My Account Create an AWS Account

Products Solutions Pricing Documentation Learn Partner Network AWS Marketplace Customer Enablement Events

Standard Reserved Instances provide you with a significant discount (up to 75%) compared to On-Demand instance pricing, and can be purchased for a 1-year or 3-year term. Customers have the flexibility to change the Availability Zone, the instance size, and networking type of their Standard Reserved Instances.

Purchase Convertible Reserved Instances if you need additional flexibility, such as the ability to use different instance families, operating systems, or tenancies over the Reserved Instance term. Convertible Reserved Instances provide you with a significant discount (up to 54%) compared to On-Demand Instances and can be purchased for a 1-year or 3-year term.

Reserved Instances Payment Options

You can choose between three payment options when you purchase a Standard or Convertible Reserved Instance. With the **All Upfront** option, you pay for the entire Reserved Instance term with one upfront payment. This option provides you with the largest discount compared to On-Demand instance pricing. With the **Partial Upfront** option, you make a low upfront payment and are then charged a discounted hourly rate for the instance for the duration of the Reserved Instance term. The **No Upfront** option does not require any upfront payment and provides a discounted hourly rate for the duration of the term.

Note: For T2 and T3 Unlimited instances, CPU Credits are charged at:

- \$0.05 per vCPU-Hour for Linux, RHEL and SLES, and
- \$0.096 per vCPU-Hour for Windows and Windows with SQL Web

The CPU Credit pricing is the same for all T2 and T3 instance sizes, for On-Demand and Reserved Instances, and across all T2 and T3 regions.

See [T2 and T3 Unlimited documentation](#) for details on when CPU Credits are charged.

Linux RHEL SLES Windows Windows with SQL Standard

Windows with SQL Web Windows with SQL Enterprise Linux with SQL Standard

Linux with SQL Web Linux with SQL Enterprise

Region: US East (Ohio)

a1.medium

STANDARD 1-YEAR TERM					
Payment Option	Upfront	Monthly*	Effective Hourly**	Savings over On-Demand	On-Demand Hourly
No Upfront	\$0	\$11.75	<u>\$0.016</u>	37%	\$0.0255
Partial Upfront	\$67	\$5.62	<u>\$0.015</u>	40%	
All Upfront	\$131	\$0.00	<u>\$0.015</u>	41%	

CONVERTIBLE 1-YEAR TERM					
Payment Option	Upfront	Monthly*	Effective Hourly**	Savings over On-Demand	On-Demand Hourly
No Upfront	\$0	\$13.51	<u>\$0.019</u>	27%	\$0.0255
Partial Upfront	\$77	\$6.42	<u>\$0.018</u>	31%	
All Upfront	\$151	\$0.00	<u>\$0.017</u>	32%	

STANDARD 3-YEAR TERM					
Payment Option	Upfront	Monthly*	Effective Hourly**	Savings over On-Demand	On-Demand Hourly
No Upfront	\$0	\$8.03	<u>\$0.011</u>	57%	

Windows Taskbar

01:35 AM 18-02-2020

Amazon EC2 Dedicated Host Pricing: Dedicated host pricing varies by instance family, regardless of the quantity or the size of the instance that you select to run. On-demand pricing gives the flexibility to scale up or down without long-term commitments.

Amazon EC2 Dedicated Host Instance Payment Option

The screenshot shows the Amazon EC2 Dedicated Host Pricing page. The page title is "Amazon EC2 Dedicated Host Pricing". The URL is "aws.amazon.com/ec2/dedicated-hosts/pricing/". The page is in English and shows the "US East (Ohio)" region. The pricing is for On-Demand Dedicated Hosts. The page lists various instance families and their corresponding prices per hour.

Region: **US East (Ohio)**

Price Per Hour

Instance Family	Price Per Hour
General purpose	
a1	\$0.449
m5	\$5.069
m5d	\$5.966
m5dn	\$7.181
m5n	\$6.283
m4	\$2.42
Compute optimized	
c5	\$3.366
c5d	\$3.802
c5n	\$4.277
c4	\$1.75
GPU instance	
p3	\$26.928
p2	\$15.84
g4dn	\$8.606
g3	\$5.016
Memory optimized	
x1	\$14.672
x1e	\$29.357
r5	\$6.653
r5d	\$7.603
r5dn	\$8.818
r5n	\$7.867
r4	\$4.682
r3	\$2.922
z1d	\$4.91
Storage optimized	
i3	\$5.491
i3en	\$11.933
i2	\$7.502
h1	\$4.118
d2	\$6.072

(References: <https://aws.amazon.com/ec2/pricing/>)

10. 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), 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. In order 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:

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

(References: <http://searchcloudcomputing.techtarget.com/essentialguide/Breaking-down-whats-in-your-cloud-SLA>, https://en.wikipedia.org/wiki/Service-level_agreement, <https://aws.amazon.com/ec2/sla/>)

b. The process what exactly one need to do to create a free machine instance, that could be used as a server are follows:

- 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/>)