[ENGLISH]

Hello,

Cloud Gurus and welcome to this lecture where we're going to explore

the various pricing options or pricing models

for EC2. So let's take a look. Now,

there are a few different pricing models with EC2.

So let's start with on demand, which is the default.

So with on demand, you pay by the hour or the second,

depending on the type of instance you're running.

So you only pay for the time when you are running your instance,

we then have reserved instances and with reserved instances,

you can reserve EC2 capacity for either one or three

years.

So you commit to using a certain amount of EC2 instances of a

specific type for either one or three years and in return for making

this commitment,

you get up to a 72% discount on the hourly charge and

reserved instances operate at a regional level.

So say for example,

you commit to using 10 instances of type T3.large in

the US East one region.

If you spin up up to 10 instances within us East one,

that will be covered by your reserved capacity but if you launch an instance in

a different region, that's not going to be covered by your reservation.

So you won't get the discount on instances that you spin up outside the region

of your reserved capacity. We then have spot instances,

and this is where you can purchase unused capacity at a massive discount of

up to 90% and the price fluctuates according to

supply and demand.

So you set a maximum price that you're willing to pay for the instance,

but the catch is that as soon as the price exceeds your maximum,

the instance will be terminated or hibernated depending on the options that

you've chosen. So you get a massive discount,

but it's not really suitable for every application and then

finally we have the dedicated option.

So this is a physical server running EC2 instances,

and it's dedicated for your use,

and it's actually the most expensive option, but it's a good option.

If you have software licenses, which are tied to physical hardware,

or if you have very strict compliance requirements,

which state that you cannot use multitenant hardware for your application.

So which option should you choose? Well, it all depends on your application.

So for on demand instances where you've got the low cost and flexibility

of EC2 without any upfront payment or longterm commitment,

this is great for applications with short term or spiky or

unpredictable workloads that cannot be interrupted and it's also really good for

testing the water.

So for applications which are being developed or tested for the very

first time,

so on demand is the most flexible option moving on to reserved

instances.

So these are great for applications with predictable usage patterns.

So applications with a steady state or very predictable usage.

They're also really good for applications which have very specific capacity

requirements and you want to be sure that the capacity is always going to be

available in the region that you want when you need it.

Reserved instances are also really good if you can pay upfront and if you make

upfront payments with your reserved instances,

then you can get the benefit of even more discounts and there are three types

of reserved instance. So first of all,

we've got the standard reserved instance where you can reserve a set quantity of

a certain instance type and they claim you can get up to a 72%

discount, but in practice,

I've never been able to get it anywhere near that much and I think it all

depends on the combination of components and options you put into the pricing

calculator and how much you're willing to pay upfront.

Now one restriction with standard reserved instances is that if you reserve

site 10 T3.large instances for one year,

and then your application requirements change,

you may need to change to a bigger instance,

but with standard reserved instances, you cannot change your mind.

So that is where convertible reserved instances come in and

with these,

you get up to 54% of the on demand price and the

great thing with these is that you get the option to change to a different

reserved instance type of equal or greater value.

So if we want to move our application from T3.large to T3.extra-large,

we can convert our reservation to any other instance,

type of equal or greater value and don't worry too much

at the moment about these instance types and instance names.

We're going to go through all of the different instance types in the next

lecture. And lastly, we have scheduled reserved instances.

So these allow you to launch a reserved instance within a time frame that

you specify and this allows you to match your capacity reservation

to a predictable recurring schedule,

which only requires either a fraction of a day, week, or month.

So this allows you to schedule your reservation for a window of time,

which suits you. So think about maybe a phone company,

calculating customer bills at the end of each month.

So they might want to use it scheduled reserved instances because they know

they're always going to use a fairly predictable,

large amount of capacity for a few days at the end of every single month

and they can use scheduled reserved instances to save some money on their AWS

bill.

So moving on to spot instances and with spot instances,

if you remember you set a maximum price that you're willing to pay for the

instance,

and then the actual price fluctuates with supply and demand,

and if the price exceeds your maximum,

then your instance can either be terminated or hibernated depending on the

options that you selected. So when would we use spot instances? Well,

they're really great for applications which have a flexible start and end time.

They also work really well for applications that are really only feasible at

very low compute prices.

So you can take advantage of the low cost of spot instances,

and they're also suitable for applications or users with an urgent

need for large amounts of additional computing capacity,

which is temporary. So it's not an ongoing requirement,

it's just for a specific job or workload.

So you might use it spot instances for workloads like image rendering or

parallel workloads like genomic sequencing or even running calculations

for algorithmic trading engines,

and then finally dedicated hosts. So when would we use it these?

Well, these are really great for meeting compliance requirements or regulatory

requirements, which may not support a multitenant architecture.

They're also great if you're in a situation where you've got software licensing,

which doesn't support multitenancy or cloud deployments,

and you can purchase them on either an on demand or reserved basis and with

on demand, they're purchased hourly and with the reserved option,

you can get up to a 70% discount off the on demand price.

Now, in addition to these pricing options,

they've also introduced something called savings plans,

which is another way to save money on your AWS bill.

So you can save up to 72% on all AWS compute usage,

regardless of instance type or region and the way it works is that you

commit to one or three years to use a specific amount of compute

power, and it's measured in dollars per hour,

rather than an instance type or size,

et cetera and it's super flexible because it doesn't just cover EC2.

It also includes serverless technologies like Lambda and Fargate.

And of course,

Lambda is the very first serverless technology allowing you to run at your

applications in the cloud without having to worry about infrastructure and

Fargate is a serverless technology allowing you to run Docker containers once

again,

without having to worry about the underlying infrastructure and then one last

thing I wanted to show you is the AWS pricing calculator and this is

a great tool for helping you understand the various pricing options within AWS.

So I recommend you take a look at the pricing calculator and explore some of the

AWS services in there,

particularly around EC2 and the various different pricing models.

So here is the pricing calculator,

and I've included a link to this in the resources section of the course.

So you just head to create estimate and you can select a service

or you can search for it.

So I'm just going to search for EC2 hit configure.

And this is where you can create a cost estimate for your EC2 instances.

So first of all, make sure your region is correct, scroll down,

and then you can select the specification of your EC2 instance,

and you can select the capacity you want either in terms of CPUs or memory,

or you can search for instances by name,

and here are all our instance types.

So I'm going to go in and select T3.medium.

I'm going to have a quantity of 10,

and this is where you can choose the pricing strategy.

So let's start off with on demand instances

and if we scroll down the page,

our total monthly cost for 10 servers is going to be

$333.

So what happens if I change that to standard reserved instances?

I'm going to reserve for one year with no upfront payments

and straight away my cost has reduced to 220.

So what if I make my reservation three years,

I get an even bigger discount. So that's over a 50% discount.

So I would definitely suggest that you spend some time playing around with the

pricing calculator and getting to know the different pricing options.

And while all these cost saving approaches are wonderful.

One of the best ways to control costs in AWS is to

remember to turn off the things that you no longer need.

So don't leave services running if you don't need them anymore and I love this

cartoon. It's a good reminder to us all,

to remember to terminate instances,

if we're no longer using them and especially if you're working in your own

AWS account,

it's really good practice to decommission everything as soon as you finished

playing with it to avoid any unnecessary costs.

So let's move on to my exam tips for EC2 pricing

options. So there's on demand,

which allows you to pay by the hour or second,

depending on the type of instance you run and this is a great option for

flexibility. We've also got reserved instances,

which allows you to reserve capacity for one or three years where the discount

they claim of up to 72% on the hourly charge. And this is great.

If you've got known fixed requirements,

we then have spot instances,

which allows you to purchase unused capacity at a discount of up to 90%

prices fluctuate with supply and demand,

and you set your maximum that you're willing to pay and if the price goes over

the maximum, they're going to either hibernate or terminate your instance,

depending on the options that you selected and this is great for applications

with flexible start and end times,

but it's not so great for applications which need to be up and running all the

time and then finally there is the dedicated option,

and this is where you get a physical EC2 server dedicated for your use

and it's really great if you've got server bound licenses that you want to

reuse, or if you have compliance requirements,

which are preventing you from using a multi-tendency.

So that is the end of this lecture. If you have any questions,

please do let me know. Otherwise,

please join me in the next lecture where we're going to take a look at all the

different EC2 instance types that are available, and there are loads of them.

So I'll see you in the next lecture. Thank you.

[SPANISH]

Hola,

Cloud Gurus y bienvenidos a esta conferencia donde vamos a explorar

las diversas opciones de precios o modelos de precios

para EC2. Así que echemos un vistazo. Ahora,

hay algunos modelos de precios diferentes con EC2.

Entonces, comencemos con a pedido, que es el predeterminado.

Así que con on demand, pagas por hora o por segundo,

según el tipo de instancia que esté ejecutando.

Por lo tanto, solo paga por el tiempo en que ejecuta su instancia,

tenemos instancias reservadas y con instancias reservadas,

puede reservar capacidad EC2 para uno o tres

años \_

Así que te comprometes a usar una cierta cantidad de instancias EC2 de un

específico durante uno o tres años y a cambio de hacer

este compromiso,

obtienes hasta un 72% de descuento en el cargo por hora y

reservadas operan a nivel regional.

Entonces digamos por ejemplo,

se compromete a usar 10 instancias de tipo T3.large en

la región este de EE.UU.

Si gira hasta 10 instancias dentro de nosotros Este,

que estará cubierto por su capacidad reservada, pero si lanza una instancia en

una región diferente, eso no estará cubierto por su reserva.

Por lo tanto, no obtendrá el descuento en las instancias que active fuera de la región.

de su capacidad reservada. Entonces tenemos instancias puntuales,

y aquí es donde puede comprar capacidad no utilizada con un gran descuento de

hasta un 90% y el precio fluctúa según

oferta y demanda

Así que establece un precio máximo que está dispuesto a pagar por la instancia,

pero el problema es que tan pronto como el precio exceda su máximo,

la instancia será terminada o hibernada dependiendo de las opciones que

has elegido. Entonces obtienes un descuento masivo,

pero no es realmente adecuado para todas las aplicaciones y luego

finalmente tenemos la opción dedicada.

Este es un servidor físico que ejecuta instancias EC2,

y está dedicado para su uso,

y en realidad es la opción más cara, pero es una buena opción.

Si tiene licencias de software, que están vinculadas al hardware físico,

o si tiene requisitos de cumplimiento muy estrictos,

que indican que no puede usar hardware multiinquilino para su aplicación.

Entonces, ¿qué opción debe elegir? Bueno, todo depende de tu aplicación.

Por lo tanto, para instancias bajo demanda en las que tiene bajo costo y flexibilidad

de EC2 sin ningún pago por adelantado o compromiso a largo plazo ,

esto es ideal para aplicaciones a corto plazo o puntiagudas o

impredecibles que no se pueden interrumpir y también es muy bueno para

probando el agua.

Así que para aplicaciones que están siendo desarrolladas o probadas para el mismo

primera vez,

así que bajo demanda es la opción más flexible pasando a reservado

instancias \_

Por lo tanto, estos son excelentes para aplicaciones con patrones de uso predecibles.

Por lo tanto, aplicaciones con un estado estable o un uso muy predecible.

También son muy buenos para aplicaciones que tienen una capacidad muy específica.

requisitos y quiere estar seguro de que la capacidad siempre va a ser

disponible en la región que desee cuando lo necesite.

Las instancias reservadas también son muy buenas si puede pagar por adelantado y si hace

por adelantado con sus instancias reservadas,

entonces puede obtener el beneficio de aún más descuentos y hay tres tipos

de instancia reservada. Entonces, en primer lugar,

tenemos la instancia reservada estándar donde puede reservar una cantidad determinada de

cierto tipo de instancia y afirman que puede obtener hasta un 72%

descuento , pero en la práctica,

Nunca he sido capaz de conseguirlo tanto y creo que todo

depende de la combinación de componentes y opciones que pongas en el precio

calculadora y cuánto está dispuesto a pagar por adelantado.

Ahora, una restricción con las instancias reservadas estándar es que si reserva

sitio 10 T3.grandes instancias durante un año,

y luego cambian los requisitos de su aplicación,

es posible que deba cambiar a una instancia más grande,

pero con las instancias reservadas estándar, no puede cambiar de opinión.

Ahí es donde entran las instancias reservadas convertibles y

con estos,

obtienes hasta el 54% del precio bajo demanda y el

mejor de estos es que tienes la opción de cambiar a un diferente

reservada de igual o mayor valor.

Entonces, si queremos mover nuestra aplicación de T3.grande a T3.extragrande,

podemos convertir nuestra reserva a cualquier otra instancia,

tipo de igual o mayor valor y no te preocupes demasiado

en este momento sobre estos tipos de instancias y nombres de instancias.

Vamos a repasar todos los diferentes tipos de instancias en el próximo

conferencia \_ Y por último, tenemos instancias reservadas programadas.

Esto le permite lanzar una instancia reservada dentro de un marco de tiempo que

especificas y esto te permite igualar tu reserva de capacidad

a un programa recurrente predecible,

que solo requiere una fracción de un día, una semana o un mes.

Entonces esto le permite programar su reserva para una ventana de tiempo,

que te conviene Así que piensa en tal vez una compañía telefónica,

calcular las facturas de los clientes al final de cada mes.

Por lo tanto, es posible que deseen usarlo en instancias reservadas programadas porque saben

van a utilizar una bastante predecible,

gran cantidad de capacidad durante unos días al final de cada mes

y pueden usar instancias reservadas programadas para ahorrar algo de dinero en su AWS

factura \_

Entonces, pasando a instancias puntuales y con instancias puntuales,

si recuerdas que estableces un precio máximo que estás dispuesto a pagar por el

instancia ,

y luego el precio real fluctúa con la oferta y la demanda,

y si el precio excede su máximo,

entonces su instancia puede ser terminada o hibernada dependiendo de la

opciones que haya seleccionado. Entonces, ¿cuándo usaríamos instancias puntuales? Bien,

son realmente geniales para aplicaciones que tienen una hora de inicio y finalización flexible.

También funcionan muy bien para aplicaciones que realmente solo son factibles en

muy bajos.

Para que pueda aprovechar el bajo costo de las instancias puntuales,

y también son adecuados para aplicaciones o usuarios con urgencia

necesidad de grandes cantidades de capacidad informática adicional,

que es temporal. Así que no es un requisito continuo,

es solo para un trabajo específico o carga de trabajo.

Por lo tanto, puede usarlo en instancias puntuales para cargas de trabajo como representación de imágenes o

paralelas como la secuenciación genómica o incluso la ejecución de cálculos

para motores de negociación algorítmica,

y finalmente hosts dedicados. Entonces, ¿cuándo lo usaríamos estos?

Bueno, estos son realmente geniales para cumplir con los requisitos de cumplimiento o reglamentarios.

requisitos , que pueden no ser compatibles con una arquitectura multiinquilino.

También son excelentes si se encuentra en una situación en la que tiene una licencia de software,

que no admite implementaciones en la nube o multiusuario ,

y puede comprarlos ya sea bajo demanda o reservados y con

bajo demanda, se compran por horas y con la opción reservada,

puede obtener hasta un 70% de descuento sobre el precio bajo demanda.

Ahora, además de estas opciones de precios,

introdujeron algo llamado planes de ahorro,

que es otra forma de ahorrar dinero en su factura de AWS.

Por lo tanto, puede ahorrar hasta un 72 % en todo el uso de computación de AWS,

independientemente del tipo de instancia o la región y la forma en que funciona es que usted

comprometerse a uno o tres años para usar una cantidad específica de cómputo

energía , y se mide en dólares por hora,

en lugar de un tipo o tamaño de instancia,

etcétera y es súper flexible porque no solo cubre EC2.

También incluye tecnologías sin servidor como Lambda y Fargate .

Y por supuesto,

Lambda es la primera tecnología sin servidor que le permite ejecutarse en su

aplicaciones en la nube sin tener que preocuparse por la infraestructura y

Fargate es una tecnología sin servidor que le permite ejecutar contenedores Docker una vez

de nuevo ,

sin tener que preocuparse por la infraestructura subyacente y luego una última

Lo que quería mostrarles es la calculadora de precios de AWS y esto es

una gran herramienta para ayudarlo a comprender las diversas opciones de precios dentro de AWS.

Así que te recomiendo que eches un vistazo a la calculadora de precios y explores algunas de las

Servicios de AWS allí,

particularmente en torno a EC2 y los diferentes modelos de precios.

Así que aquí está la calculadora de precios,

y he incluido un enlace a esto en la sección de recursos del curso.

Así que simplemente diríjase a crear una estimación y puede seleccionar un servicio

o puedes buscarlo.

Así que solo voy a buscar EC2 hit configure.

Y aquí es donde puede crear una estimación de costos para sus instancias EC2.

Antes que nada, asegúrese de que su región sea correcta, desplácese hacia abajo,

y luego puede seleccionar la especificación de su instancia EC2,

y puede seleccionar la capacidad que desee en términos de CPU o memoria,

o puede buscar instancias por nombre,

y aquí están todos nuestros tipos de instancias.

Así que voy a entrar y seleccionar T3.medium.

Voy a tener una cantidad de 10,

y aquí es donde puede elegir la estrategia de precios.

Entonces, comencemos con instancias bajo demanda

y si nos desplazamos hacia abajo en la página,

nuestro costo mensual total para 10 servidores será

$333.

Entonces, ¿qué sucede si cambio eso a instancias reservadas estándar?

Voy a reservar por un año sin pagos por adelantado

y de inmediato mi costo se ha reducido a 220.

¿Y qué si hago mi reserva tres años,

Tengo un descuento aún mayor. Eso es más del 50% de descuento.

Así que definitivamente te sugiero que pases un tiempo jugando con el

de precios y conocer las diferentes opciones de precios.

Y aunque todos estos enfoques de ahorro de costos son maravillosos.

Una de las mejores formas de controlar los costos en AWS es

recuerda apagar las cosas que ya no necesitas.

Así que no dejes los servicios abiertos si ya no los necesitas y me encanta esto.

dibujos animados Es un buen recordatorio para todos nosotros,

recordar terminar las instancias,

si ya no los usamos y especialmente si estás trabajando por tu cuenta

cuenta AWS,

es una muy buena práctica desmantelar todo tan pronto como termine

jugando con él para evitar costos innecesarios.

Así que pasemos a mis consejos de examen para los precios de EC2

opciones \_ Así que hay bajo demanda,

que te permite pagar por hora o por segundo,

dependiendo del tipo de instancia que ejecute y esta es una gran opción para

flexibilidad \_ También tenemos instancias reservadas ,

que permite reservar capacidad para uno o tres años donde el descuento

reclaman hasta un 72% sobre la tarifa horaria. Y esto es genial.

Si tiene requisitos fijos conocidos,

tenemos instancias puntuales,

que le permite comprar capacidad no utilizada con un descuento de hasta el 90%

los precios fluctúan con la oferta y la demanda,

y estableces el máximo que estás dispuesto a pagar y si el precio supera

como máximo, hibernarán o terminarán su instancia,

dependiendo de las opciones que haya seleccionado y esto es ideal para aplicaciones

con horas de inicio y finalización flexibles,

pero no es tan bueno para las aplicaciones que necesitan estar en funcionamiento todo el tiempo

tiempo y finalmente está la opción dedicada,

y aquí es donde obtiene un servidor EC2 físico dedicado para su uso

y es realmente genial si tiene licencias vinculadas al servidor que desea

reutilizar , o si tiene requisitos de cumplimiento,

que le impiden utilizar una multitendencia.

Así que ese es el final de esta conferencia. Si tienes alguna pregunta,

por favor hágamelo saber. De lo contrario,

por favor acompáñenme en la próxima conferencia donde vamos a echar un vistazo a todos los

diferentes tipos de instancias EC2 que están disponibles, y hay muchos de ellos.

Así que te veré en la próxima conferencia. Gracias.