# Prework Sesión 2

# Introducción

AWS es un entorno complejo, hay docenas de servicios que pueden ser manejados con diferentes herramientas, pero antes de hacer eso es necesario comprender la forma en que AWS organiza su propia infraestructura y en consecuencia sus servicios, conocer de entre todos los servicios cuales son los servicios principales (spoiler: son principales por que otros servicios usan estos servicios principales) y por que pueden representar ventajas con respecto a otros modelos de infraestructura.

# 1. Objetivo 🎯

* Conocer la interfaz de AWS Console (portal web).
* Comprender las formas de interacción con los servicios de AWS.
* Explorar la taxonomía de la infraestructura global de AWS.
* Conocer los servicios principales de AWS.
* Quedar claro con las ventajas que representa consumir servicios en la nube.
* Evitar sorpresas desagradables en la factura de fin de mes, si se tiene una sorpresa brindar las herramientas necesarias para detectar el problema.

# 2. Instrucciones 📋

* Se revisarán los conceptos básicos sobre la infraestructura de AWS, atender los enlaces externos para formar mas contexto al respecto.

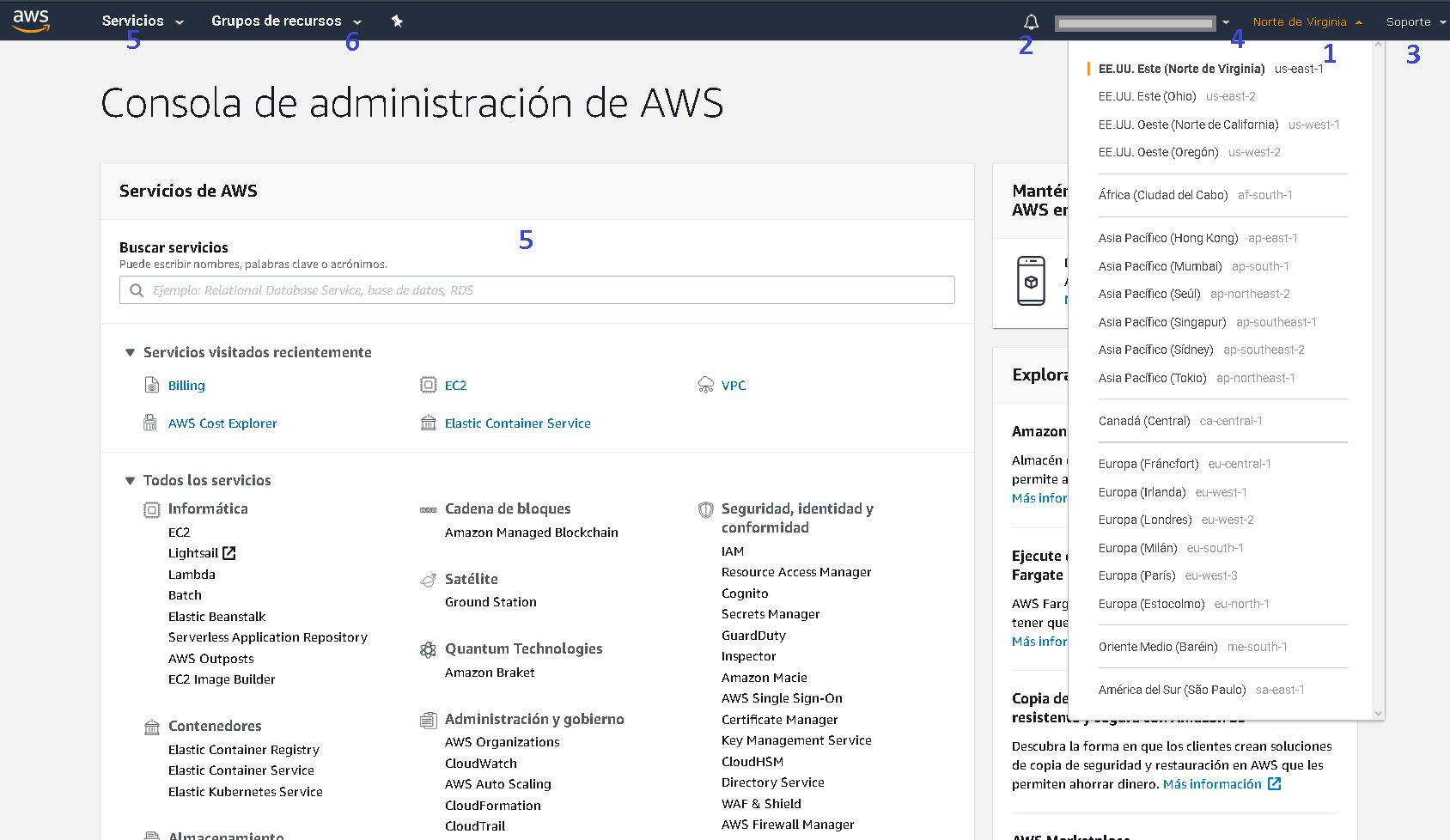
# 3. Desarrollo 📑

## ¿Por qué AWS?

Según [Gartner](https://www.gartner.com/reviews/market/public-cloud-iaas/vendor/amazon-web-services/product/amazon-web-services/ratings?reviewRating=3&sort=-helpfulness), el rubro de integración y despliegue es marcado con un 4.6/5 de excelencia, en general el despliegue toma menos de tres meses, lo que es un tiempo razonable (aunque hay que decir que no se menciona el tamaño de las organizaciones), en temas de soporte que es muy importante también se encuentra con un 4.6/5, las organizaciones encuentran muy bueno el tiempo de respuesta y la calidad del soporte recibido es muy buena. En cuanto a servicios la calificación va de 4.7/5 con el puntero en servicios de cómputo (EC2), user management y networking, también seguridad y apego a estándares son muy bien calificados.

# Interactuando con AWS y configuración de cuentas

Al ingresar la cuenta de AWS lo primero que veremos es lo siguiente:



interactuando-con-la-consola

1.- Muchos de los servicios de AWS operan en regiones específicas, desde este menú se selecciona la región para poner en marcha servicios de AWS, ¿qué tal desplegar la página web de la organización en Londres?

2.-Las alertas brindan información sobre algunas fallas o degradamiento de servicios que puede presentar AWS, es especialmente importante cuando el usuario final reporta problemas con el software ejecutándose sobre AWS, se pueden ver las alarmas y verificar si uno de los servicios de la aplicación esta implicado en problemas.

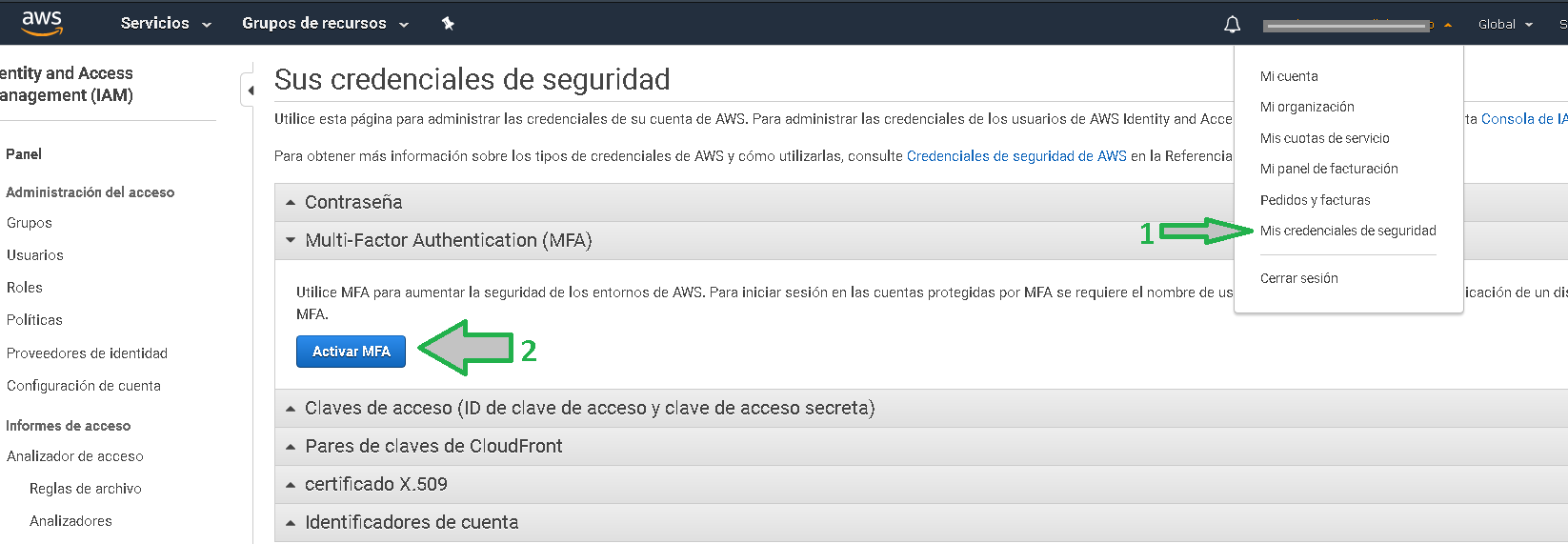
3.- En el menú soporte se pueden dar de alta tickets de soporte y dependiendo del plan de soporte va desde reportar problemas y aclaraciones con la facturación hasta reportar caídas de los aplicativos. Aqui se pide también a AWS aumentar el límite de los recursos de la cuenta. El soporte esta disponible el Japonés e Inglés vía web o telefónica.

4.- En el menú de usuario se pueden administrar las tarjetas para hacer los pagos, asociar otras cuentas para administrarlas, ver las facturas y establecer políticas de seguridad en cuanto a la cuenta se refiere.

5.- El menú de servicios se puede navegar entre todos los servicios de AWS disponibles en cualquier momento.

6.- El menú grupo de recursos contiene herramientas que ayudarán a organizar los recursos de AWS que utilizados. Con los grupos de recursos se agrupan servicios en la misma región, también se pueden poner etiquetas a los recursos que usados a fin de utilizar como parámetros de filtrado en otras secciones como el Explorador de costos.

Lo más importante y lo primero a realizar a tener una cuenta de AWS es habilitar el doble factor de autenticación, con ello se disminuye la probabilidad drásticamente que alguien use la cuenta de una manera no autorizada o en el peor de los casos secuestren la cuenta con todos los recursos, datos y aplicaciones que se estén ejecutando en ella. Si es una cuenta personal tal vez las consecuencias sean mínimas, pero si es una cuenta de un cliente donde tiene toda su empresa se puede volver un problema muy grave si alguien no autorizado gana acceso a la cuenta. Para habilitarlo habrá que ir al menú de usuario, dar click en “Mis credenciales” y de ahí click en el botón “Habilitar MFA“, la manera mas común es usar alguna aplicación como Google Authenticator, solo que hay que tomar en cuenta que si se daña el celular o se pierde se corre el riesgo a ya no tener acceso a la cuenta de AWS, una alternativa es la app [Authy](authy.com), que puede ser configurada en múltiples dispositivos incluidas laptops y PC de escritorio con lo que se reduce la probabilidad de dejar de tener acceso, y claro, siempre se puede acceder con una clave secreta mostrada al momento de dar de alta el factor múltiple de autenticación, esta clave se debe guardar ante la falta de acceso a los dispositivos con el token.



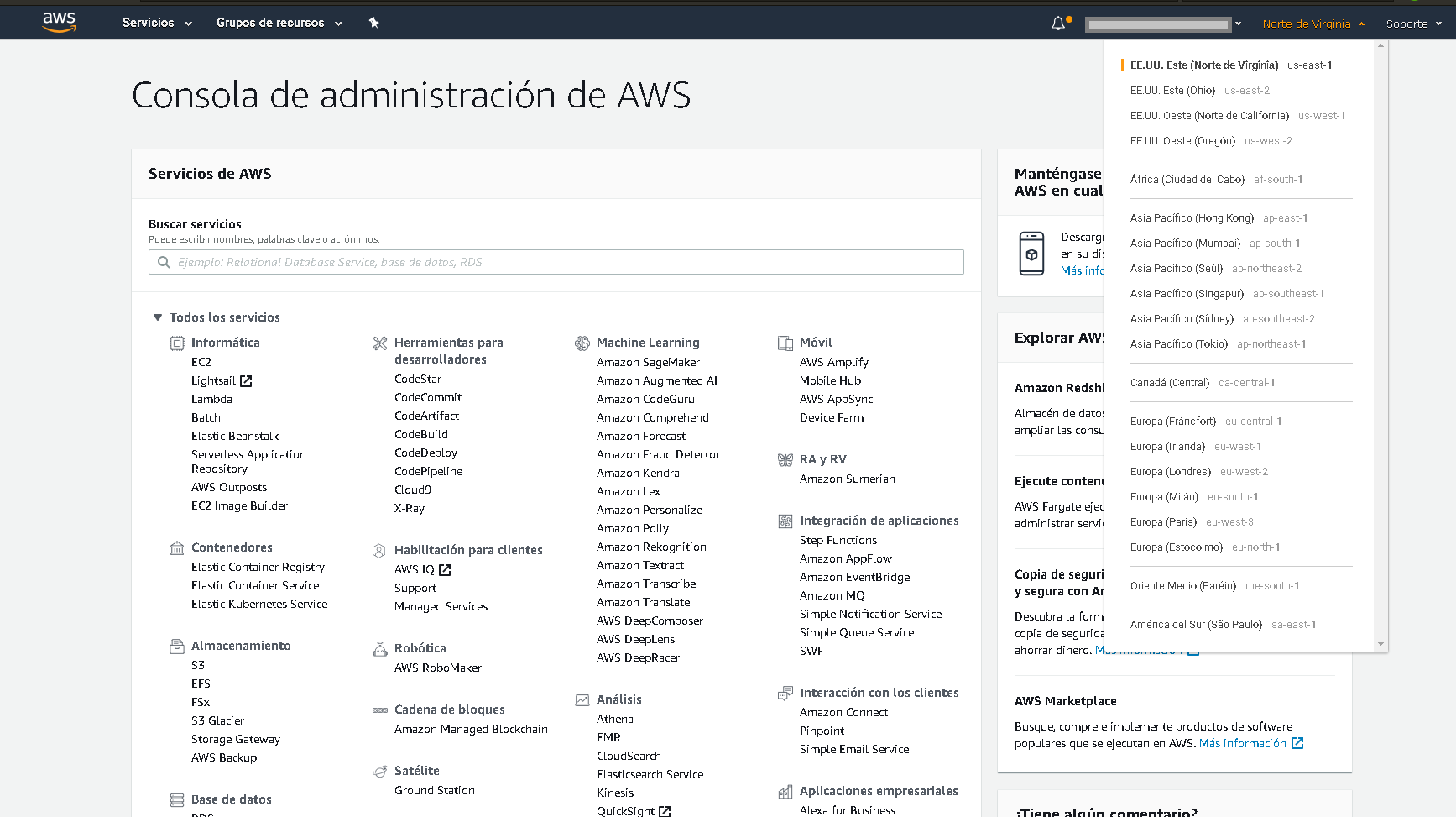
multifactor.png

# AWS Console, AWS CLI & AWS SDK

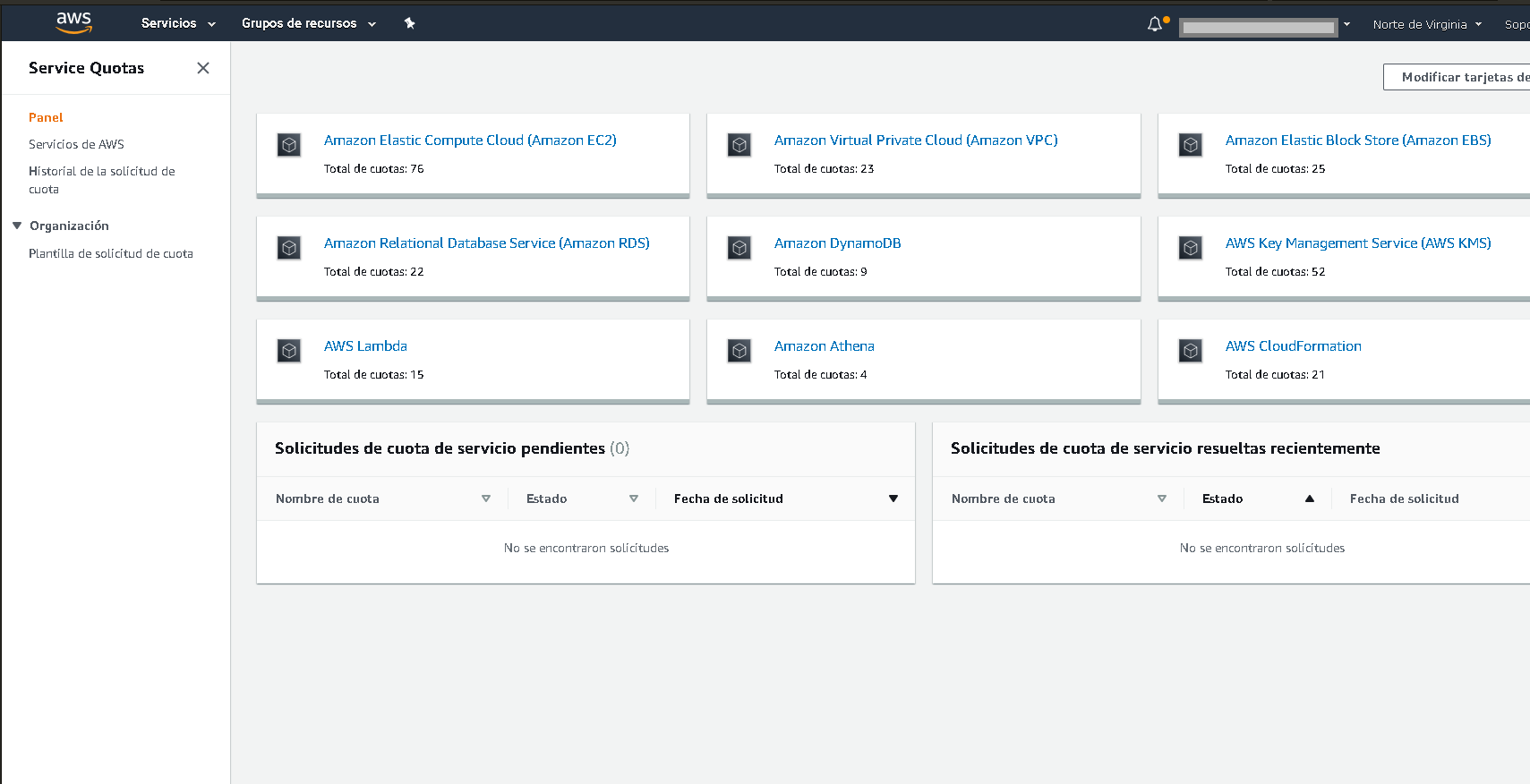
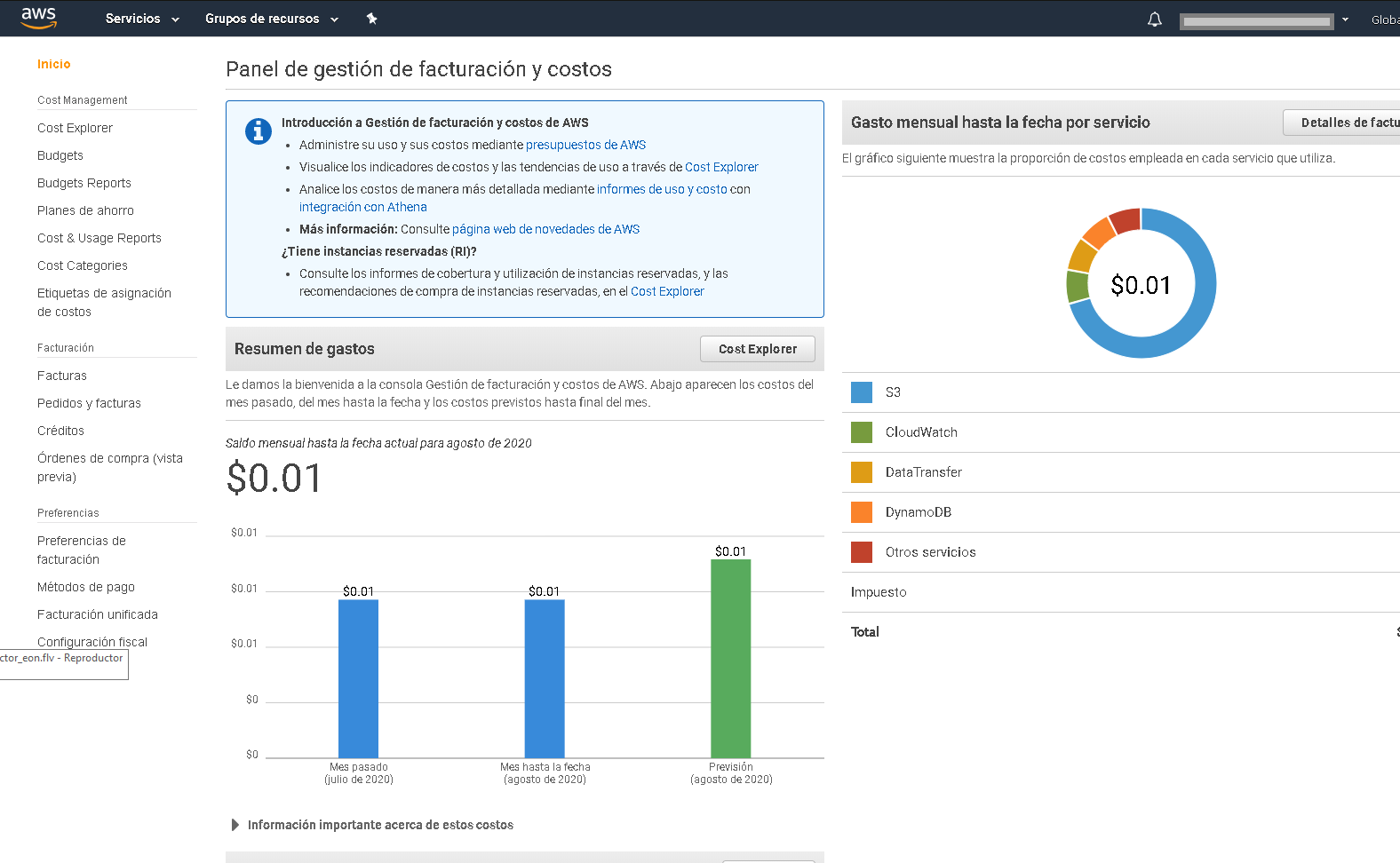
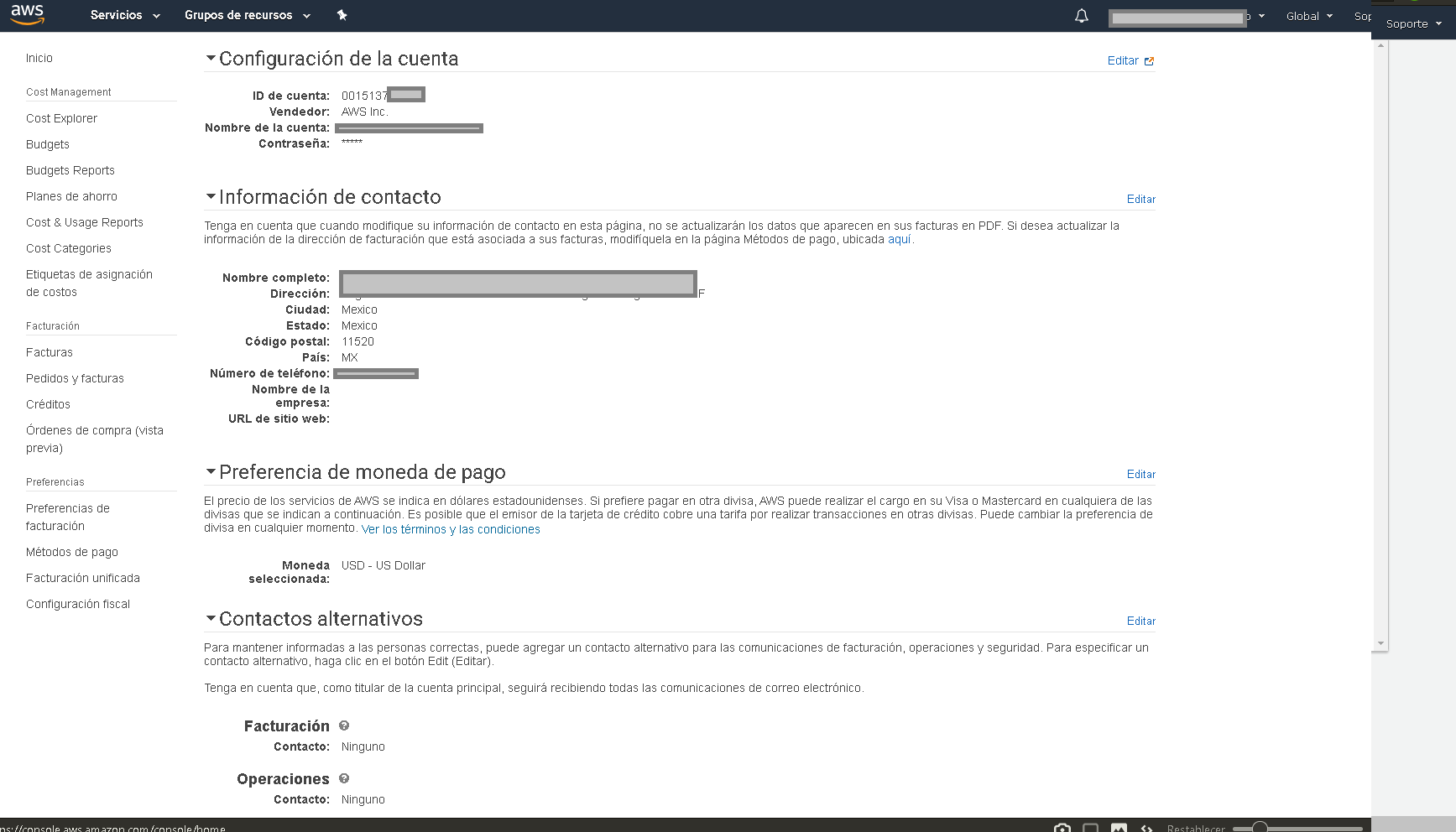
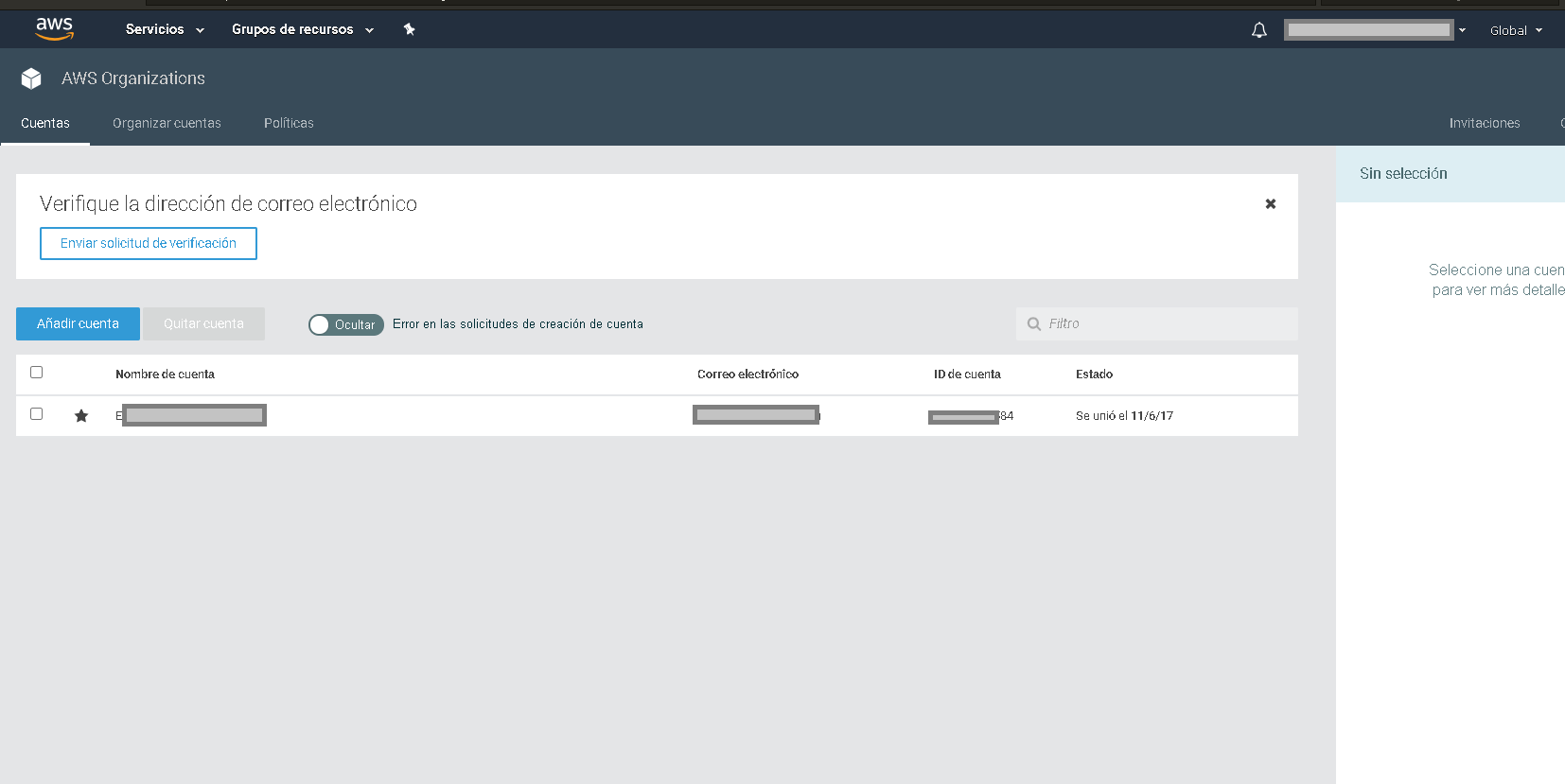
En general se puede interactuar con los recursos y servicios que brinda AWS de 3 modos distintos.

## AWS Console:

Para la gente visual, esta es la forma mas natural para tener un primer acercamiento con AWS, al entrar al [portal web](console.aws.amazon.com/) de la consola de AWS se pueden ver y gestionar los diferentes recursos que ofrece AWS.

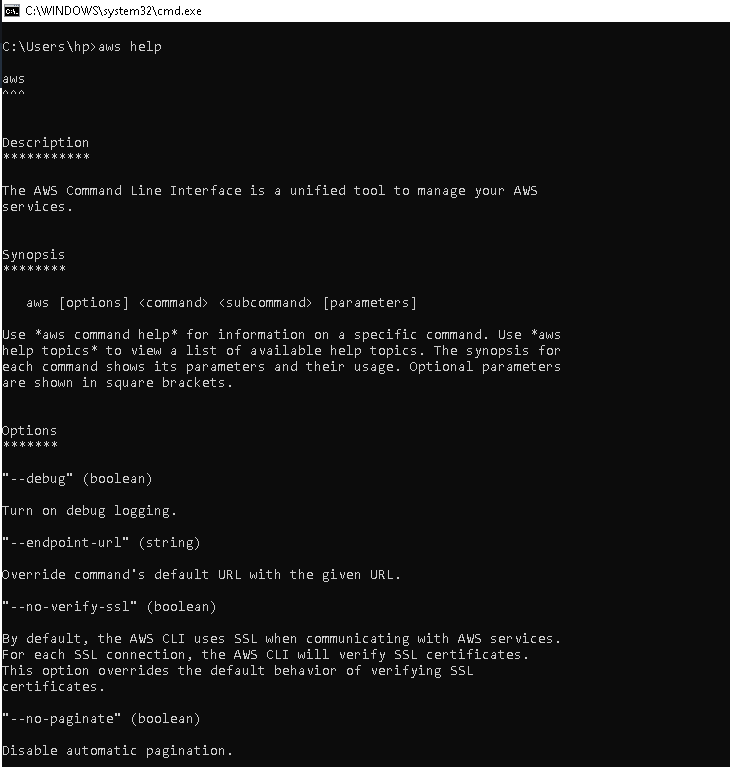
Esta es la apariencia que tiene la consola de administración de AWS hasta Agosto del 2020. 

El elemento principal de la consola es el buscador de servicios que ofrece, sin el buscador es fácil perderse entre los 175 que se ofrecen. Desde aquí se pueden manejar:

* Cuotas: Hay servicios de AWS que tienen límites predefinidos, por ejemplo con las instancias de máquinas virtuales (EC2) son [estas](https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-on-demand-instances.html#ec2-on-demand-instances-limits). , desde esta sección se puede apelar a aumentar las cuotas de los servicios si la operación de la organización así lo requiere. Es importante señalar que hay cuotas de tipo soft que pueden ser cambiadas solicitando a AWS un aumento y de tipo hard que no pueden ser cambiadas. [Por ejemplo](https://docs.aws.amazon.com/mediapackage/latest/ug/live-quotas.html).
* Facturación: Tener visibilidad de los costos, presupuestos y métodos de pago es muy importante para la correcta gestión y estrategia en los activos de TI hospedados en AWS. . Aquí hay mucha ayuda visual sobre la cantidad de dinero gastada mes a mes, se puede entrar al detalle de la facturación para saber cuánto dinero es gastado por cada recurso de AWS, se pueden establecer presupuestos y alarmas si es que dichos presupuestos se rebasan también es posible configurar el envió de reportes con detalle para personas de la organización en intervalos definidos.
* Configuración de la cuenta: Se gestionan los datos propios de la cuenta como nombre del propietario de la cuenta, teléfono y dirección al igual la gestión de contraseña. 
* Organizaciones: AWS provee los mecanismos necesarios para el manejo de múltiples cuentas centralizadamente. . Por ejemplo, se pueden manejar desde aquí las cuentas de los clientes si se provee algún servicio de consultoría, así los clientes no tienen que dar acceso de su cuenta principal al consultor.

## AWS CLI

También existe la opción de manejar los recursos de AWS por medio de línea de comandos disponible para los sistemas [operativos mas comunes](https://docs.aws.amazon.com/cli/latest/userguide/install-cliv2.html), es cómodo usar la línea de comandos ante tareas repetitivas que exijan automatización, se deben descargar archivos a diario desde S3?, es posible generar un script de linea de comandos que lo haga por el administrador, así no hay que preocuparse por entrar a la consola web que tiende a ralentizar el proceso.



aws-cli

# AWS SDK

Para los programadores también hay opción de manejar los servicios de AWS desde su [lenguaje de programación](https://aws.amazon.com/es/tools/) de preferencia, Ruby, Python, PHP, Node.js, .NET, JavaScript, C++, Go y Java son los lenguajes soportados. ¿Cómo se puede usar?, se puede generar una aplicación web que en lugar de guardar las imágenes de las identificaciones de los usuarios en el disco duro del servidor las guarde directamente en el servicio de S3 de AWS.

# Infraestructura Global de AWS (AWS Regions, Availability Zones, AWS Edge Locations, and Local Zones)

El concepto de nube puede sonar un tanto abstracto, pero al final no hay que dejar de recordar que son un conjunto de máquinas agrupadas en un edificio en algún lugar del planeta.

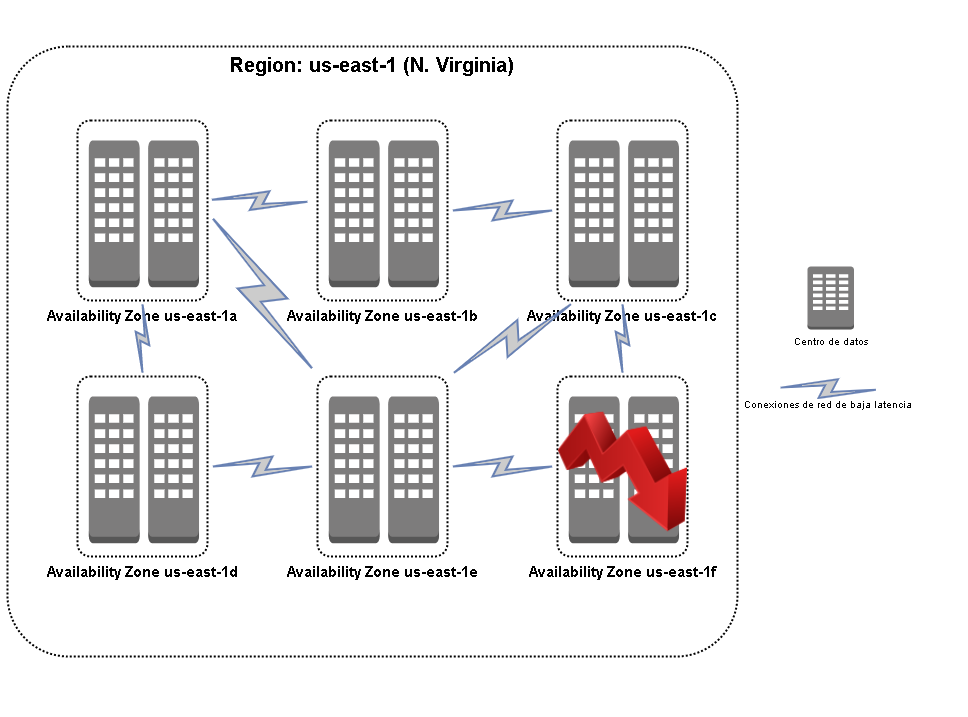
Amazon Web Services tiene centros de datos lleno de servidores por todo el planeta buscando brindar a los usuarios disponibilidad, rendimiento, escalabilidad, y flexibilidad.

Los conceptos claves que ayudarán a comprender como AWS tiene distribuidos todos sus servidores por el mundo son: - Zonas de disponibilidad: Los distintos centros de datos de Amazon por el mundo los agrupa las llamadas zonas de disponibilidad. Cada zona de disponibilidad puede constar de mas de un centro de datos, cada zona de disponibilidad se comunica con otras zonas de disponibilidad por medio de enlaces de baja latencia. La razón de segmentar una zona de disponibilidad en varios centros de datos con algunos kilómetros de distancia es ser resilientes ante desastres naturales, fallas de energía eléctrica, atentados o algún otro inconveniente, mas adelante se verá como maximizar la tolerancia a fallos de los aplicativos si se alojan en mas de una zona de disponibilidad.

* Regiones: Son puntos geográficos distantes donde se agrupan varias zonas de disponibilidad. AWS tiene actualmente 245 países con 24 regiones lanzadas, anunciadas 3 por lanzar en Indonesia, Japón y España.
* Local Zones: Aún esta en una etapa temprana de desarrollo, solo se encuentra una local zone disponible en Los Angeles California, pretende proveer servicios de baja latencia a usuarios finales acercando la infraestructura de AWS sin necesidad de tener una zona de disponibilidad cercana.
* Wavelength: Ofrecerá acceso óptimo en latencia a recursos de AWS por medio de redes 5G. Aún es un servicio muy cerrado a los Estados Unidos de América.
* Edge locations: Son centros de datos manejados por partners certificados de AWS que tiene conexión directa a la infraestructura de AWS. Es usado por los servicios como Route53 (servicio de DNS), CloudFront (Content Delivery Network) y API Gateway (creación, publicación y mantenimiento de API). La idea es proporcionar baja latencia a los usuarios finales no importando si se encuentran cerca de una región de AWS.

Ejemplo de las zonas de disponibilidad de la región 1, en ella se aprecian las zonas de disponibilidad (conunto de centros de datos) conectadas por redes de baja latencia. Así si la zona de disponibilidad us-east-1f tiene bajo performance por fallas por energía eléctrica las aplicaciones pueden seguir operando si están distribuidas en más de una zona de disponibilidad. Algunos servicios como RDS obligan al [provisionar](www.redhat.com/en/topics/automation/what-is-provisioning) una base de datos establecer la red en dos zonas de disponibilidad como mínimo.

### Nota: En general los servicios de AWS son muy resistentes a fallas, hay que evaluar bien si la aplicación a ejecutar justifica el uso de recursos replicados en más de una zona de disponibilidad, ya que los recursos replicados aunque no se usen al final cuestan prácticamente lo mismo que los recursos que si se están utilizando, dicho de otro modo, la infraestructura total cuesta aproximadamente el doble.



aws-regions-availability-zone

La mayoría de los servicios que ofrece AWS son inherentes a una región específica, si se desea provisionar un servidor virtual (EC2) en China habrá que desplazarse en la consola AWS (o especificar en aws-cli o aws-sdk) a la región deseada. Aunque hay algunas excepciones, servicios como IAM (gestión de usuarios, políticas de acceso y permisos), Cloud CloudFront (Content Ddelivery Network) y S3 (Almacenamiento masivo de archivos) se consideran que no pertenecen a ninguna región específica, son considerados servicios globales.

# Servicios Centrales

Los siguientes servicios son centrales para tener una unidad de procesamiento computacional operando en AWS:

* IAM: Es uno de los pilares de la seguridad en AWS ya que administra el acceso a servicios y recursos de AWS. Prácticamente es un servicio omnipresente, estará en cualquier servicio u aplicación a desplegar en la nube no importanto si es muy pequeña o muy grande. IAM permite definir permisos a niveles muy granulares lo que da mucha flexibilidad a la hora de definir quien puede acceder y que puede hacer en los recursos usados. IAM soporta la estrategia de autorización Attribute-based access control (ABAC) por medio de TAGS en los recursos, con lo que se pueden definir políticas de acceso o denegacion a un servicio basadas solo en las etiquetas de los recursos. También soporta la estrategia de autorización role-based access control (RBAC). Se pueden definir permisos (policies) en este caso a los usuarios mismos que que se crean en AWS. Siempre se debe garantizar que las cuentas que acceden a los recursos cumplan con en principio de mínimo privilegio.
* VPC (Virtual Private Cloud): Los recursos de AWS deben comunicarse unos con otros. Las maquinas virtuales ejecutadas por el servicio EC2, las funciones Lambda, las bases de datos de RDS son solo algunos ejemplos de servicios que requieren especificar una conexión de red para poder funcionar. El espacio donde se pueden configurar esas redes se conoce como VPC, que no es mas que una entidad lógica que contiene distintos componentes de red en una región dada. Una región puede tener configuradas varias VPC, se puede configurar una VPC para cada departamento de una organización, así se tendría una VPC para todo el departamento financiero y una para todo el departamento de IT, ¿qué se gana?, aislamiento, por defecto los recursos contenidos dentro de una VPC no pueden comunicarse, aunque si hay forma de hacerlo, es posible establecer un [AWS PrivateLink](https://aws.amazon.com/es/privatelink/) que en conjunto con un router de AWS puede redireccionar ese tráfico a subredes muy específicas, si ese tráfico se desea asegurar más se pueden configurar las ACL (Network access control list) a una o varias subredes configurando el tipo de tráfico que será permitido para entrar y salir, si aún eso no es suficiente se puede configurar un firewall adicional que opera a nivel de máquina virtual EC2 llamado security group. Todo eso funciona cuando es tráfico interno, pero cuando llegue el momento de comunicar servicios desde y hacia internet se tiene que agregar a la VPC los llamados NAT gateways y los Internet Gateways, el primero hace posible comunicar recursos que se encuentran en una subred que no debe tener acceso directo a internet como pueden ser redes dedicadas de base de datos, de hecho la forma para hacer una subred privada es que no tenga forma de routear tráfico hacia un Internet Gateway, Por otro lado un Internet Gateway es el mecanismo que da acceso desde internet hacia los recursos mapeando 1:1 una IP pública [(Elastic IP)](https://docs.aws.amazon.com/vpc/latest/userguide/vpc-eips.html), el tráfico hacia internet también es gestionado por Internet Gateway. Solo puede haber un único Internet Gateway por VPC, AWS garantiza que no se convertirá en un cuello de botella o un punto único de falla ya que tras bambalinas Internet Gateway es redundante y escalable.

La forma en que se organizan las redes en AWS no es muy diferente a como se hacen en un centro de datos tradicional, en AWS existe el concepto análogo de subred sobre la cual se define el espacio de direcciones IP [(CDIR)](https://whatismyipaddress.com/cidr) que la red puede solventar. Las subredes aquí tienen la particularidad que están sociadas a una y solo una zona de disponibilidad, al generar las subredes habrá que tener cuidado donde son creadas y que niveles de alta disponibilidad se desea y claro, el presupuesto con el que se cuenta ya que los datos entre zonas de disponibilidad tienen costo pro GB en ambas direcciones. Para la redirección y manejo del tráfico se cuentan en AWS con Routers, los routers definen las reglas o rutas que determinan a donde será dirigido el tráfico. Se pueden comunicar las subredes desde y hacia internet con NAT Gateway e Internet Gateway. VPC se puede conectar a las instalaciones de la organización por VPN con un Virtual Private Gateway, habilitando así el acceso a la red híbrida conectando infraestructura propia con la nube brindando lo mejor de los dos mundos.

* S3 (Simple Storage Service):

S3 es uno de los servicios icónicos de AWS. S3 lanzado en 2006 se ha convertido en el storage por excelencia a la hora de guardar archivos (objetos) a precios bajos con tiempos de acceso de milisegundos, el hecho de ser altamente resistente a pérdida de archivos ofreciendo un 99,999999999 % de durabilidad de datos (si se almacenan 10 000 objetos se espera una pérdida de un objeto cada 10 000 000 años) brinda mucha seguridad y confianza. S3 cuenta con diferentes niveles de servicio y precios, hay tipos de storage mas baratos pero el tiempo de acceso a los datos es mayor, si se requiere acceso a datos en el orden de los milisegundos el storage tiende a ser mas costoso. La forma de trabajar con S3 es por medio de buckets, los buckets son un contenedor de archivos u objetos, a cada bucket se le puede asignar un nombre único que no puede ser usado en ninguna otra región ni en ninguna otra cuenta de AWS, los buckets también pueden ser configurados con políticas de seguridad, aquí también es donde se especifica el nivel de storage a requerir y en consecuencia los costos que incurrirán. Los casos de uso comunes para S3 incluyen backups por su porcentaje de durabilidad, datalakes por su capacidad de escalamiento de unos cuantos gigabytes a petabytes sin esfuerzo, como storage para datos no utilizados que se deben guardar por regulaciones y como web server.

Un uso interesante además de guardar las facturas de clientes y proveedores que por reglamentación oficial en México se deben guardar tenerlas almacenadas por 5 años a bajo costo. Otra posibilidad de habilitar un bucket de S3 como un servidor web estático, hay algunas consideraciones, para poder darle un dominio a ese bucket el DNS de del dominio debe ser manejado por el servicio Route53, y para poder habilitar el certificado https se hace uso de AWS CloudFront y AWS Certificate Manager.

* EC2 y EBS Elastic Compute Cloud, o EC2, es uno de los servicios de AWS enfocados en proveer una capa de cómputo a las aplicaciones. EC2 en particular permite la creación de máquinas virtuales, también conocidas como instancias, a las cuales se les asigna una capacidad de cómputo específica en términos de la cantidad y tipos de CPU, la cantidad de memoria RAM y el espacio en disco para almacenamiento.

Adicionalmente, EC2 permite seleccionar el sistema operativo que se desea instalar en la **instancias** y con esto **entrega** una máquina virtual completamente funcional en cuestión de segundos, sobre las cuales los usuarios podrán desplegar aplicaciones y servicios de la misma manera que lo harían en un servidor on-premise o en una máquina virtual desplegada en cualquier otro hipervisor.

Las instancias EC2 requieren de una conexión de red en AWS para funcionar. Aquí es donde se debe tener previamente configurada una VPC con las subredes, grupos de seguridad, routing y niveles de disponibilidad necesarios antes de crear una maquina virtual.

Otro elemento fundamental en las instancias ec2 es el storage donde los datos serán alojados. Este storage es conocido como Elastic Block Storage. Amazon EBS proporciona volúmenes de almacenamiento duraderos a nivel de bloque que pueden ser adjuntados a una instancia en ejecución. Es posible utilizar Amazon EBS como dispositivo de almacenamiento principal para los datos que requieren actualizaciones frecuentes y granulares. Por ejemplo, Amazon EBS es la opción de almacenamiento recomendada cuando ejecuta una base de datos en una instancia.

Un EBS se comporta como un dispositivo de bloque externo **sin formato y sin formato que puede adjuntar** a una sola instancia. El volumen persiste independientemente de la vida útil de una instancia. Una vez que se adjunta un volumen de EBS a una instancia, puede usarse como cualquier otro disco duro físico.

AWS cuenta con varios tipos de instancias dependiendo de las necesidades específicas de la aplicación, se tienen:

|  |  |  |
| --- | --- | --- |
| Tipo de instancia | ¿para que usar? | Casos de uso |
| A1 | Fin general | Servidores web, Microservicios en contenedores |
| C4 | Optimizada para computación | Flotas frontend de alto rendimiento, servidores web, procesamiento por lotes |
| C5 | Optimizada para computación | Deducciones de Machine learning y otras aplicaciones que requieran un uso intensivo de recursos informático |
| C5a | Optimizada para computación | Usos que requieren un ancho de banda alto de CPU virtual y memoria, como el procesamiento por lotes, el análisis distribuido, las transformaciones de datos |
| C5ad | Optimizada para computación | Usos que requieren un ancho de banda alto de CPU virtual y memoria, como el procesamiento por lotes, el análisis distribuido, las transformaciones de datos |
| C5d | Optimizada para computación | Usos que requieren un ancho de banda alto de CPU virtual y memoria, como el procesamiento por lotes, el análisis distribuido, las transformaciones de datos |
| C5n | Optimizada para computación | Deducciones de Machine learning y otras aplicaciones que requieran un uso intensivo de recursos informático |
| D2 | Optimizada para almacenamiento | Informática distribuida de MapReduce y Hadoop |
| F1 | Computación acelerada | Acelerar algoritmos que usan muchos recursos informáticos |
| G3 | Computación acelerada | Utilizan GPU Tesla M60 de NVIDIA |
| G4 | Computación acelerada | Utilizan las GPU NVIDIA Tesla |
| H1 | Optimizada para almacenamiento | Cargas de trabajo basadas en MapReduce, sistemas de archivos distribuidos como HDFS y MapR-FS, procesamiento de datos o registros como Apache Kafka |
| I3 | Optimizada para almacenamiento | Bases de datos NoSQL |
| l3en | Optimizada para almacenamiento | Bases de datos transaccionales escalables, sistemas de archivos distribuidos, almacenamiento de datos, Elasticsearch |
| M4 | Fin general | Bases de datos de tamaño pequeño y mediano |
| M5 | Fin general | Bases de datos pequeñas y medianas, Servidores de backend para SAP, Microsoft SharePoint, otras aplicaciones empresariales |
| M5a | Fin general | Bases de datos pequeñas y medianas, Servidores de backend para SAP, Microsoft SharePoint, otras aplicaciones empresariales |
| M5ad | Fin general | Ejecución de servidores backend para SAP, Microsoft SharePoint. |
| M5d | Fin general | Ejecución de servidores backend para SAP, Microsoft SharePoint. |
| M5dn | Fin general | Ejecución de servidores backend para SAP, Microsoft SharePoint. |
| M5n | Fin general | Servidores web y de aplicaciones, bases de datos pequeñas y medianas |
| P2 | Computación acelerada | Utilizan GPU de NVIDIA Tesla K80 |
| P3 | Computación acelerada | Utilizan GPU de NVIDIA Tesla V100 |
| P3dn | Computación acelerada | Utilizan GPU de NVIDIA Tesla V100 |
| R4 | Optimizada para memoria | Bases de datos de alto rendimiento, extracción y el análisis de datos. |
| R5 | Optimizada para memoria | Bases de datos de alto rendimiento, las cachés en memoria distribuidas de escala web. |
| R5a | Optimizada para memoria | Bases de datos en memoria de tamaño mediano, los análisis de big data en tiempo real |
| R5ad | Optimizada para memoria | Bases de datos en memoria de tamaño mediano, los análisis de big data en tiempo real |
| R5d | Optimizada para memoria | Bases de datos en memoria de tamaño mediano, los análisis de big data en tiempo real |
| R5dn | Optimizada para memoria | Bases de datos en memoria de tamaño mediano, los análisis de big data en tiempo real |
| R5n | Optimizada para memoria | Bases de datos en memoria de tamaño mediano, los análisis de big data en tiempo real |
| T2 | Fin general | Sitios web y aplicaciones web,Entornos de desarrollo, compilación, pruebas e implementación |
| T3 | Fin general | Sitios web y aplicaciones web,Entornos de desarrollo, compilación, pruebas e implementación |
| T3a | Fin general | Sitios web y aplicaciones web,Entornos de desarrollo, compilación, pruebas e implementación |
| X1 | Optimizada para memoria | Bases de datos en memoria (p. ej.: SAP HANA), motores de procesamiento de big data (p. ej.: Apache Spark o Presto), informática de alto rendimiento (HPC). |
| X1e | Optimizada para memoria | Bases de datos de alto rendimiento, bases de datos en memoria (por ejemplo, SAP HANA) y aplicaciones con uso intensivo de memoria |
| z1d | Optimizada para memoria | Ideal para la automatización del diseño electrónico (EDA) |

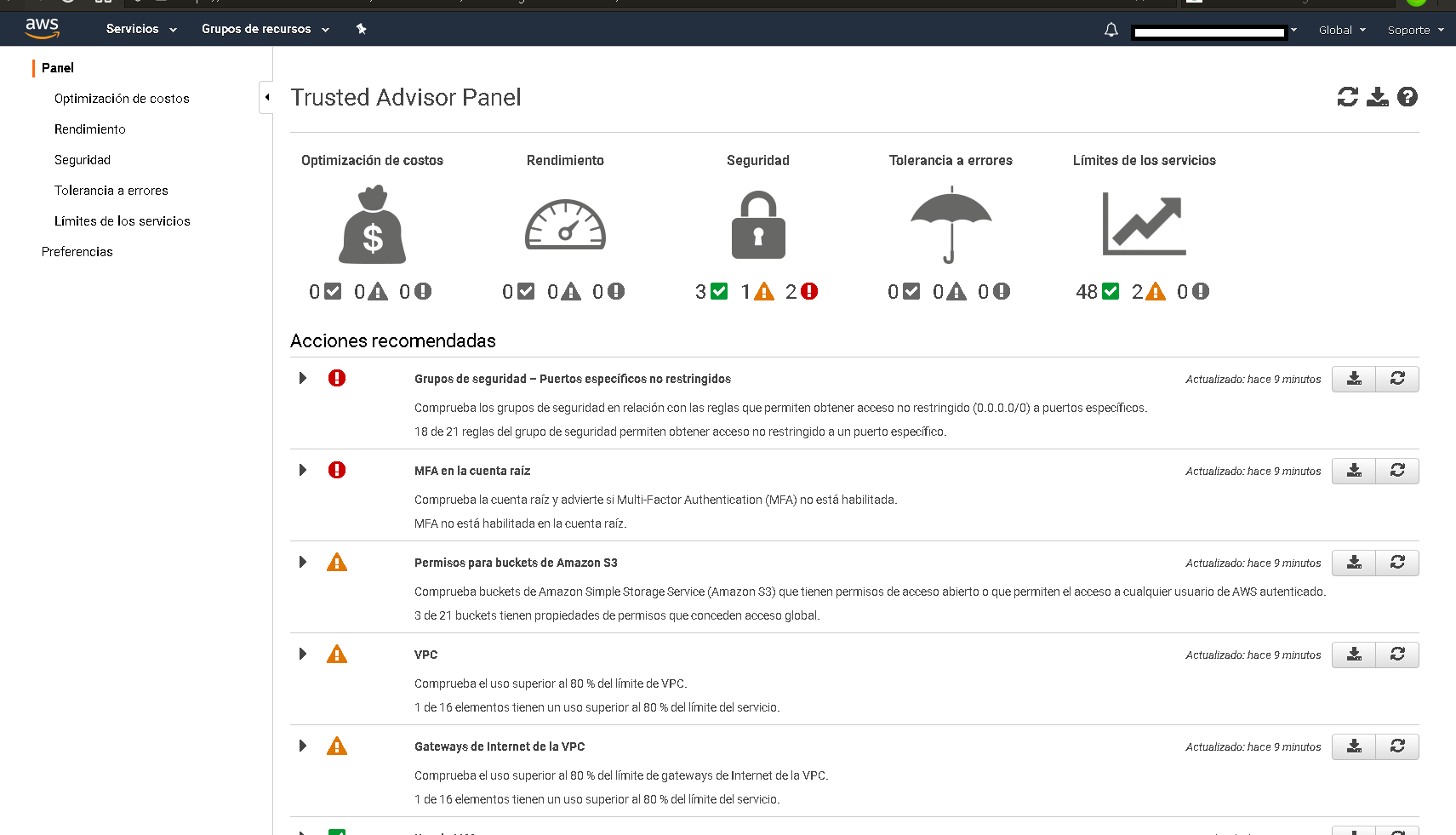
Importante mencionar que es una Amazon Machine Image (AMI). Una AMI es una máquina virtual pre configurada, usada para crear una máquina virtual.

# Niveles y planes de AWS Support

El hecho de no contar con servidores físicos y evitar todo el mantenimiento que conlleva no quiere decir necesariamente que no se requiera de algún respaldo o soporte en un momento dado. La nube de AWS cuenta con 175 servicios al momento, desde storage, redes, bases de datos, analítica, despliegue y desarrollo por mencionar algunas, por lo que no es difícil perderse en algún momento.

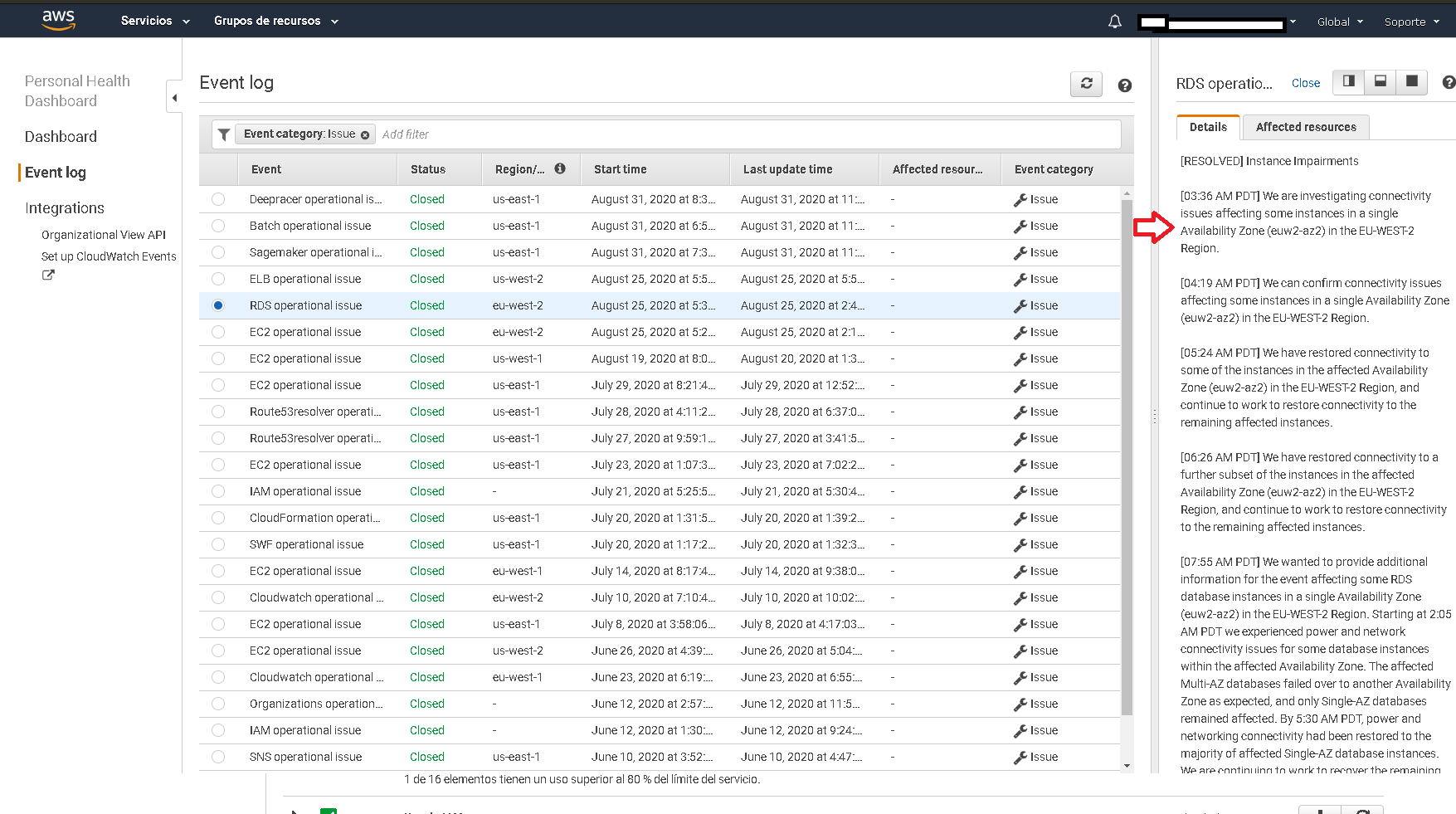
Todas las cuentas desde que son creadas cuentan con un plan Basic Support, se puede tener un soporte básico en la [documentacion](https://docs.aws.amazon.com/), y [foros de soporte](https://forums.aws.amazon.com/index.jspa).

También se cuenta con la herramienta Trust Advisor que brinda algunas recomendaciones y alertas en la cuenta de AWS con limitantes o restricciones si no se cuenta con un plan más avanzado.



trust-advisor

También se cuenta con la herramienta [Personal Health Dashboard](https://aws.amazon.com/es/premiumsupport/technology/personal-health-dashboard/) que avisa si hay problemas con los servicios de AWS y en que medida.

En la siguiente imagen se observan varios problemas que potencialmente pueden afectar a una aplicación dependiente de servicios AWS, ejemplo: si se usa un balanceador de carga y se reporta por parte de AWS en este panel que el servicio de balanceo de carga (Elastic Load Balancer) tiene pérdida de paquetes, no habrá mucho que podamos hacer pero por lo menos da visibilidad de que puede estar fallando antes de comenzar la búsqueda y pruebas en el aplicativo. 

Los siguientes son los niveles de soporte ofrecidos por AWS, cada uno da mas respaldo que el anterior, son: Developer, Bussiness y Enterprise. Si el aplicativo es relacionado con la industria espacial, sistemas de aguas de una ciudad, sistema de un hospital tal vez sean necesarios tiempos de asesoramiento técnico de menos de 4 horas. Asesoría por parte de un experto en cuanto a arquitectura por mencionar algo tal vez el plan más idóneo para este caso es Enterprise, aunque hay que tomar en cuenta que no es un servicio asequible, ya que se tendrá que estar consumiendo por lo menos 150 000 USD o pagar 15 000 USD al mes para poder acceder a él.

Si se tiene una aplicación más modesta donde es tolerable una falla de algunas horas no crítica para la organización se puede optar por el nivel Bussiness, si se facturan por lo menos 100 dólares al mes ya se puede obtener.

### NOTA: no es difícil alcanzar 100 dólares de cuenta al final de mes, una instancia de RDS de base de datos, un servidor EC2 con 100 GB de storage y un bucket con algunas decenas de gigas rondan alrededor de 67 USD de gastos al mes. Esta aplicación no es una aplicación de alta demanda, es un portal web para alta y baja de personal.

A continuación se detalla.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Desarrollador | Business | Enterprise |
|  | Se recomienda si esta realizando pruebas en AWS. | Recomendado si tiene cargas de producción en AWS. | Recomendado si tiene cargas de trabajo críticas o empresariales en AWS. |
| Comprobaciones de prácticas recomendadas de AWS Trusted Advisor | 7 comprobaciones principales | Conjunto completo de comprobaciones | Conjunto completo de comprobaciones |
| Soporte técnico mejorado | Acceso por email a los socios de soporte en la nube durante el horario comercial\*\* Casos ilimitados, 1 contacto principal | Acceso por teléfono, email y chat a los socios de soporte en la nube las 24 horas y los 7 días de la semana Casos ilimitados, contactos ilimitados (compatible con IAM) | Acceso por teléfono, email y chat a los socios de soporte en la nube las 24 horas y los 7 días de la semana Casos ilimitados, contactos ilimitados (compatible con IAM) |
| Gravedad de los casos / plazos de respuesta\* | Asesoramiento general: < 24 horas\*\* Fallo en el sistema: < 12 horas\*\* | Asesoramiento general: < 24 horas Fallo en el sistema: < 12 horas Fallo en el sistema de producción: < 4 horas Sistema de producción inactivo: < 1 hora | Asesoramiento general: < 24 horas Fallo en el sistema: < 12 horas Fallo en el sistema de producción: < 4 horas Sistema de producción inactivo: < 1 hora Sistema crítico para la empresa inactivo: < 15 minutos |
| Asesoramiento sobre arquitecturas | Aspectos generales | Contextualizado a sus casos de uso | Asesoramiento y revisión consultiva en función de sus aplicaciones |
| Administración de casos de programación |  | API de AWS Support | API de AWS Support |
| Soporte para software de terceros |  | Interoperabilidad, orientación sobre configuración y solución de problemas | Interoperabilidad, orientación sobre configuración y solución de problemas |
| Programas proactivos |  | Acceso a Infrastructure Event Management por una tarifa adicional | Infrastructure Event Management Revisiones de Well-Architected Revisiones de operaciones El director técnico de cuenta (TAM) coordina el acceso a los programas y otros expertos de AWS según sea necesario |
| Administración técnica de cuentas |  |  | Director técnico de cuenta (TAM) asignado para que monitoree de forma proactiva el entorno y apoye con los asuntos relacionados con la optimización |
| Capacitación |  |  | Acceso a laboratorios autoguiados online |
| Asistencia para cuentas |  |  | Equipo de soporte Concierge |
| Precios | A partir de 29 USD al mes\*\*\* o bien 3 % del consumo mensual de AWS       Consulte los detalles sobre precios y un ejemplo. | A partir de 100 USD al mes\*\*\* o bien 10 % del consumo mensual dACe AWS para los primeros 0 a 10 000 USD 7 % del consumo mensual de AWS desde 10 000 hasta 80 000 USD 5 % del consumo mensual de AWS desde 80 000 hasta 250 000 USD 3 % del consumo mensual de AWS por encima de 250 000 USD Consulte los detalles sobre precios y un ejemplo. | A partir de 15 000 USD o bien, 10 % del consumo mensual de AWS para los primeros 0 a 150 000 USD 7 % del consumo mensual de AWS desde 150 000 hasta 500 000 USD 5 % del consumo mensual de AWS desde 500 000 hasta 1 000 000 USD 3 % del consumo mensual de AWS por encima de 1 000 000 USD Consulte los detalles sobre precios y un ejemplo. |

# Cloud Economics

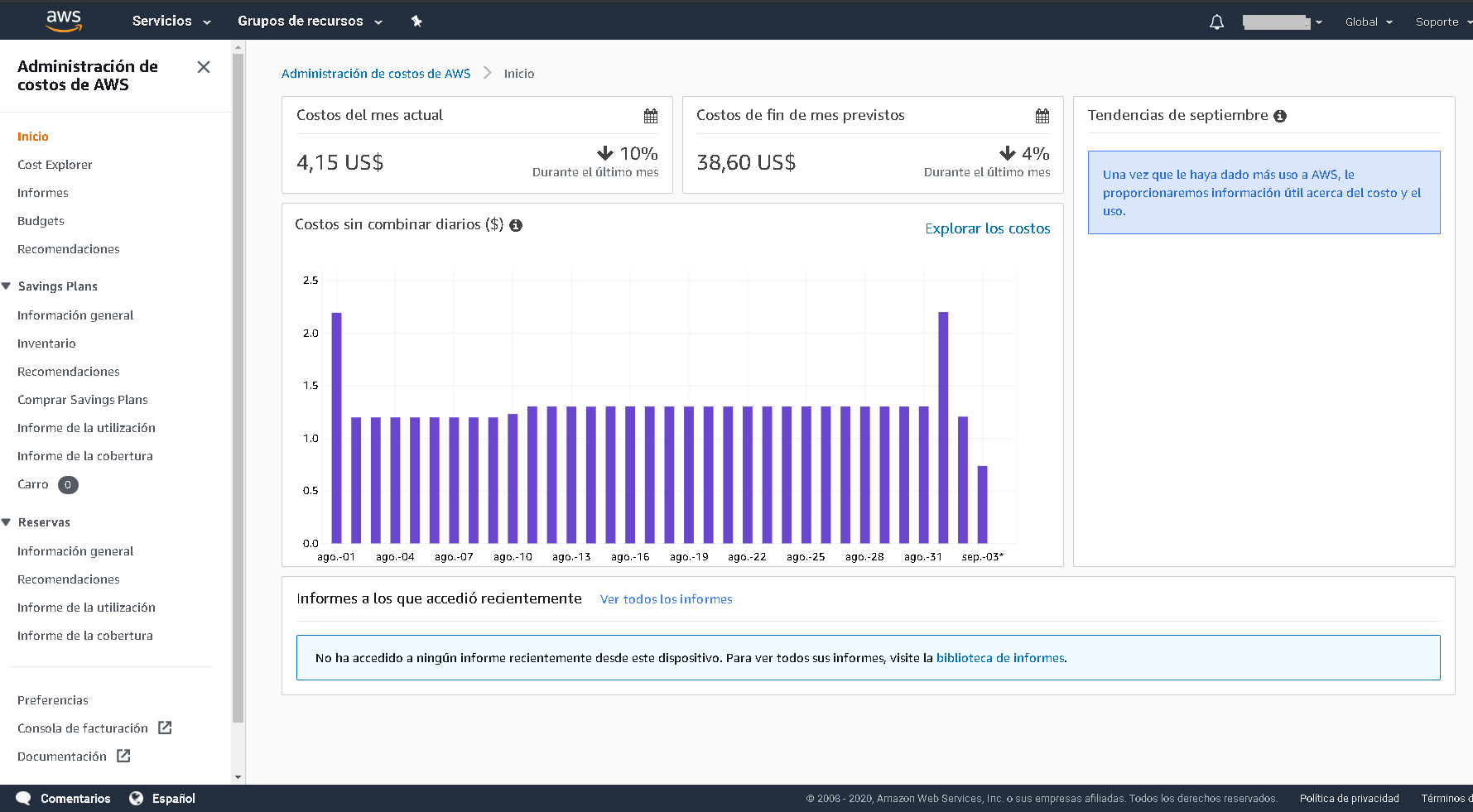
Hasta se ha navegado por algunos de los principales servicios de AWS, algunas conclusiones son:

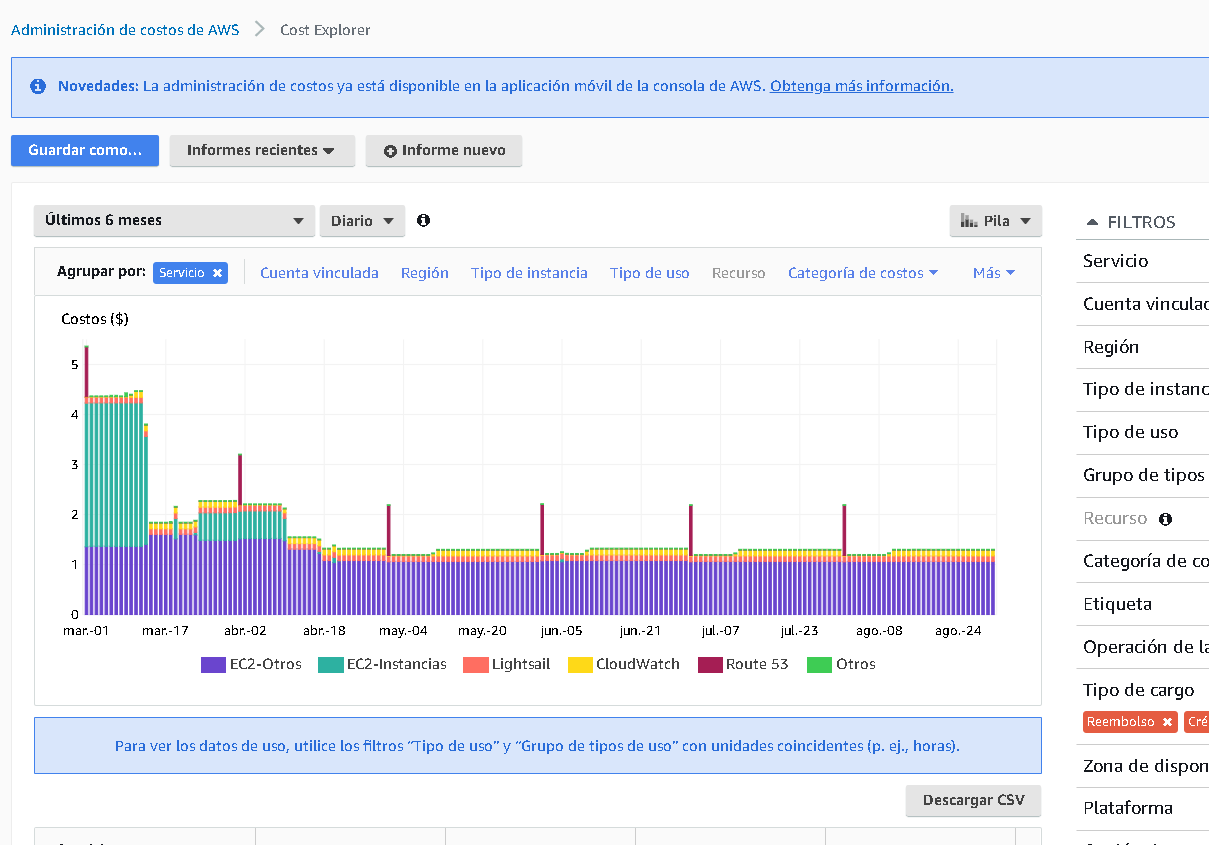
* El cómputo en la nube es accesible desde la pequeña y mediana empresa hasta multinacionales con presupuestos millonarios, tener la capacidad de contratar cómputo para mejorar y ser innovador en procesos en menos de un día (en cómputo en sitio podía llevar semanas o meses) es un antes y un después en el uso de recursos computacionales.
* Flexibilidad: Pagamos por lo que usado, se usa lo que se requiere en un tiempo determinado, hoy la carga de trabajo puede ser alta y mañana no, los recursos se pueden ajustar a estas fluctuaciones, además, se cuenta con instancias muy pequeñas desde 1 CPU y 512MB de RAM hasta instancias de 128 CPU y 2 TB de memoria, añadido a la oferta de instancias con GPU.
* En general el cómputo en la nube con el tiempo tiende a bajar los precios (aunque son centavos a lo largo de varios servicios usando es cuando el ahorro es representativo), AWS a requerir un volumen considerable de suministros como discos, energía memorias, servidores y redes puede obtener mejores precios en el mercado lo que se traduce en el consumidor en precios relativamente bajos.
* Alta disponibilidad: Se puede acceder a recursos de AWS que se encuentran en varias regiones por todo el mundo, se pueden construir aplicaciones tolerantes a fallas sin salir del país.

Indudablemente el cómputo en la nube ayudará a generar resultados más rápidos, adaptándose a la realidad y necesidades que el software tiene hoy día.

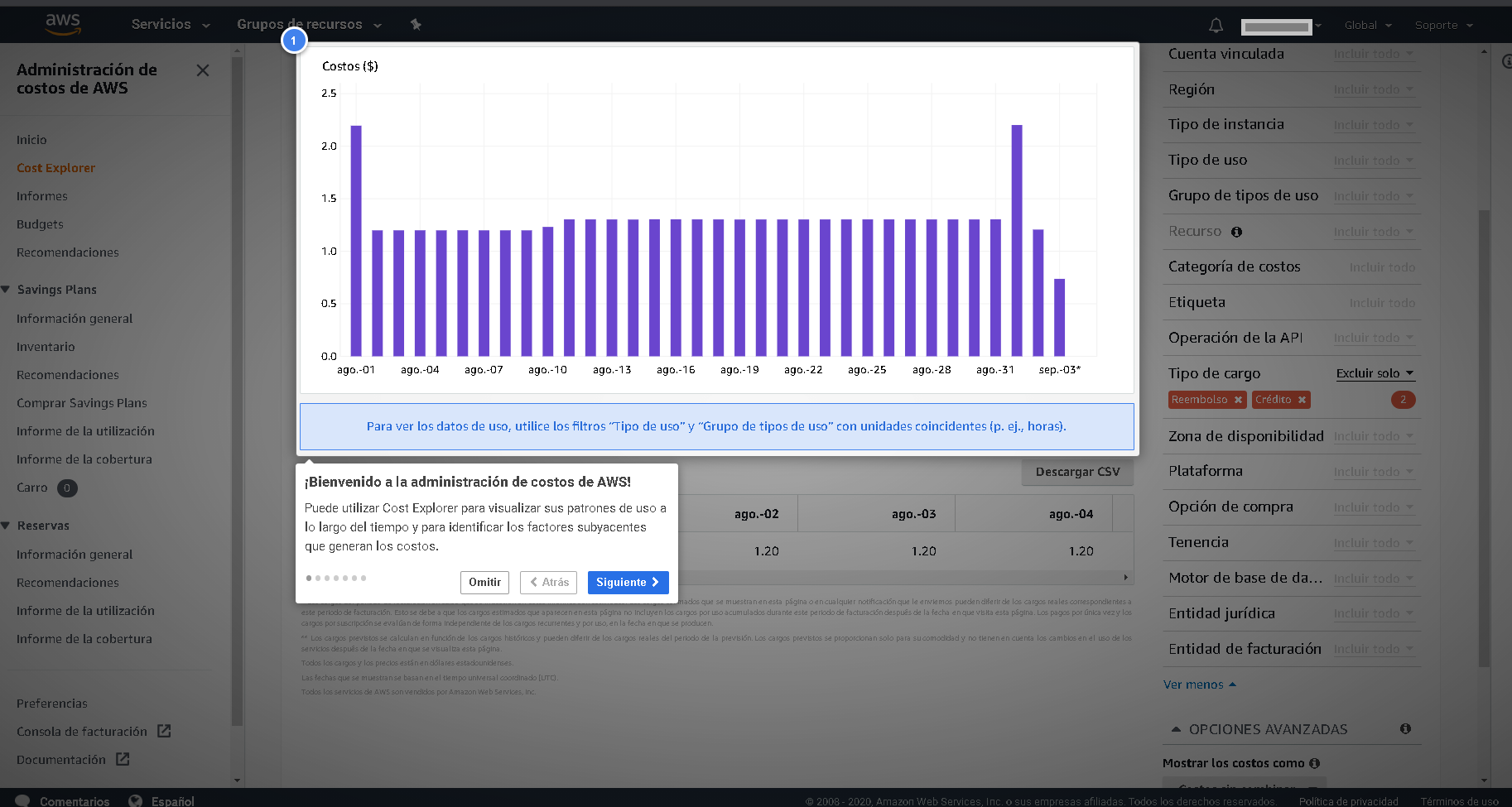
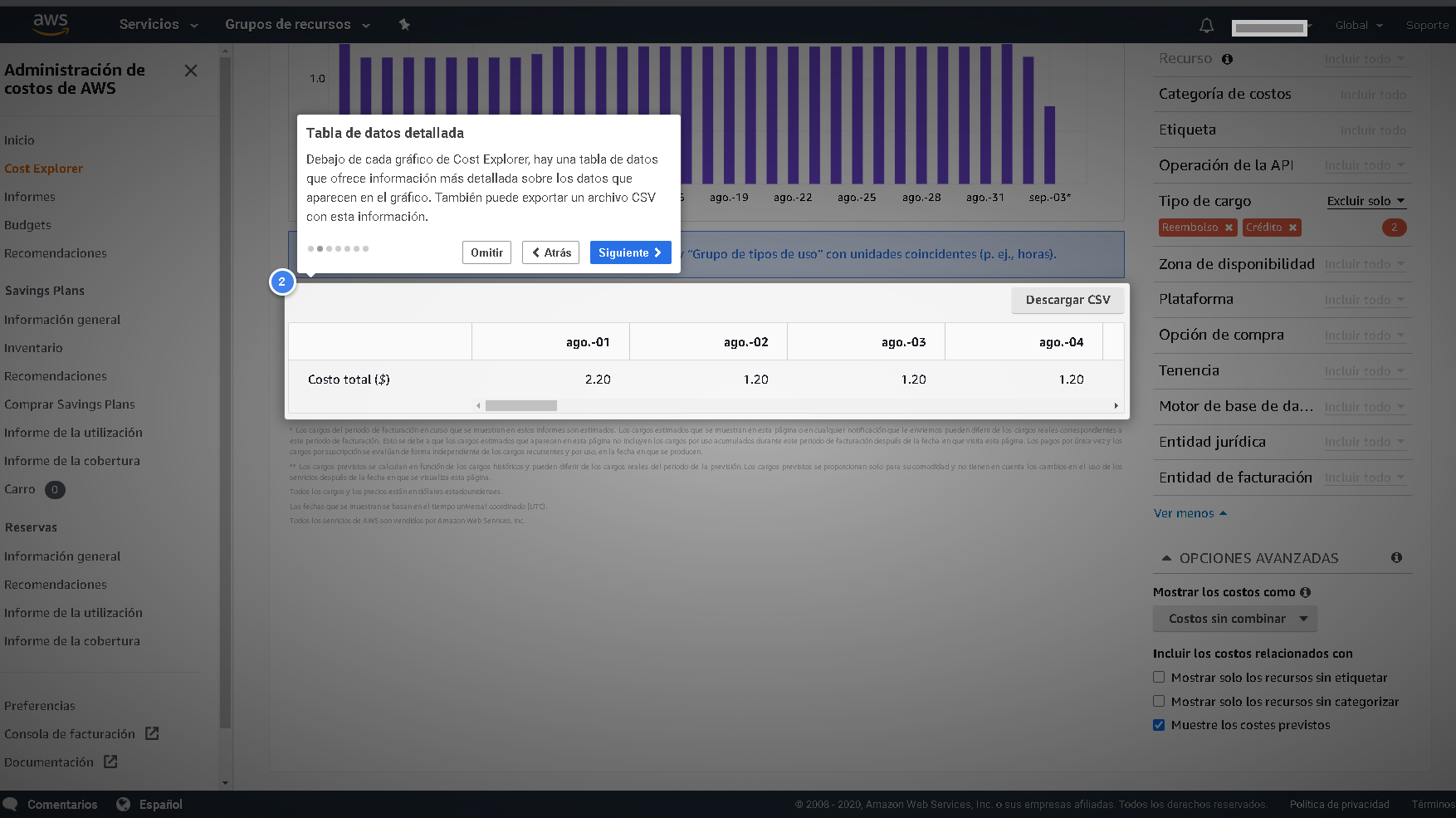
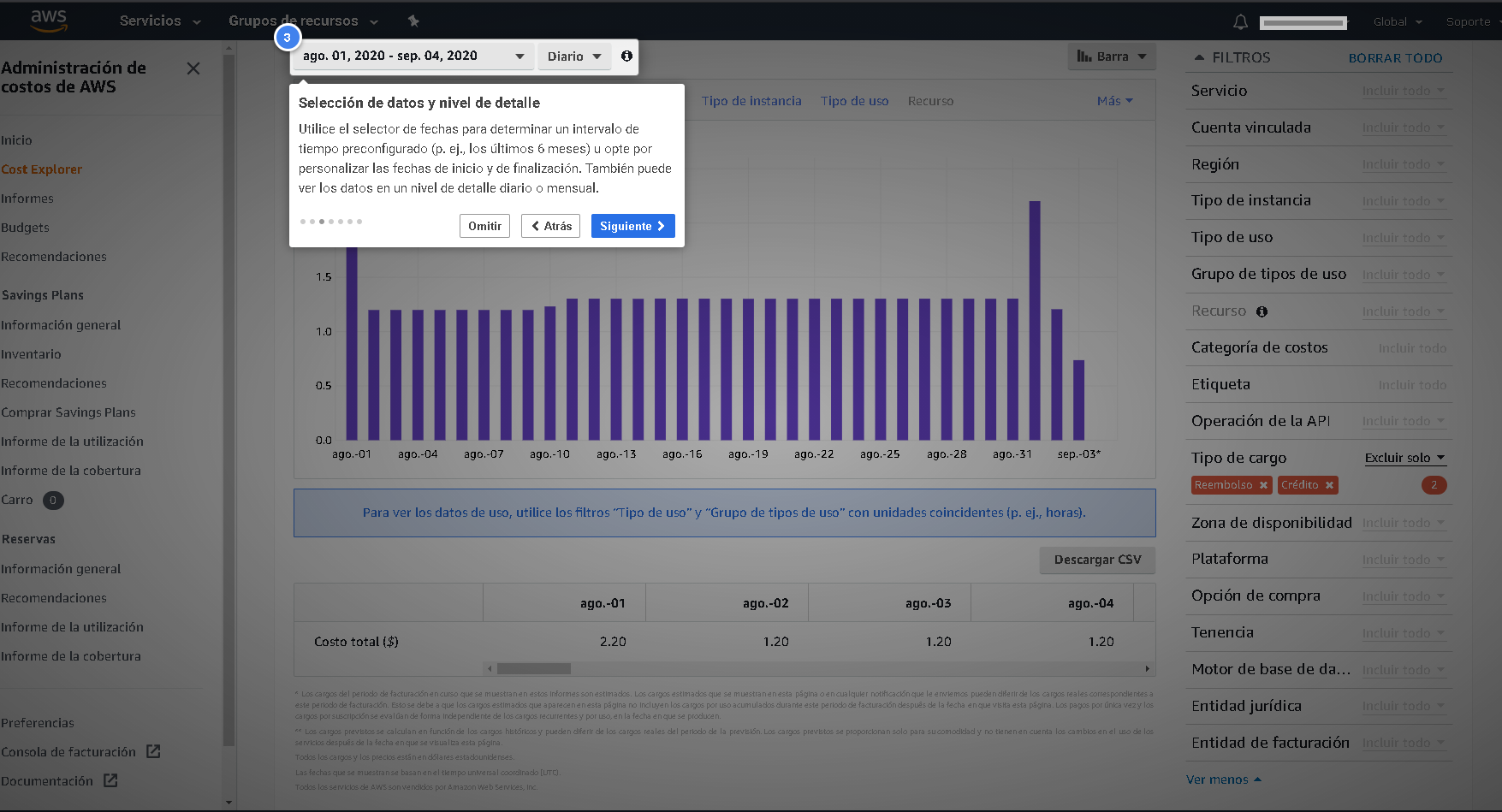
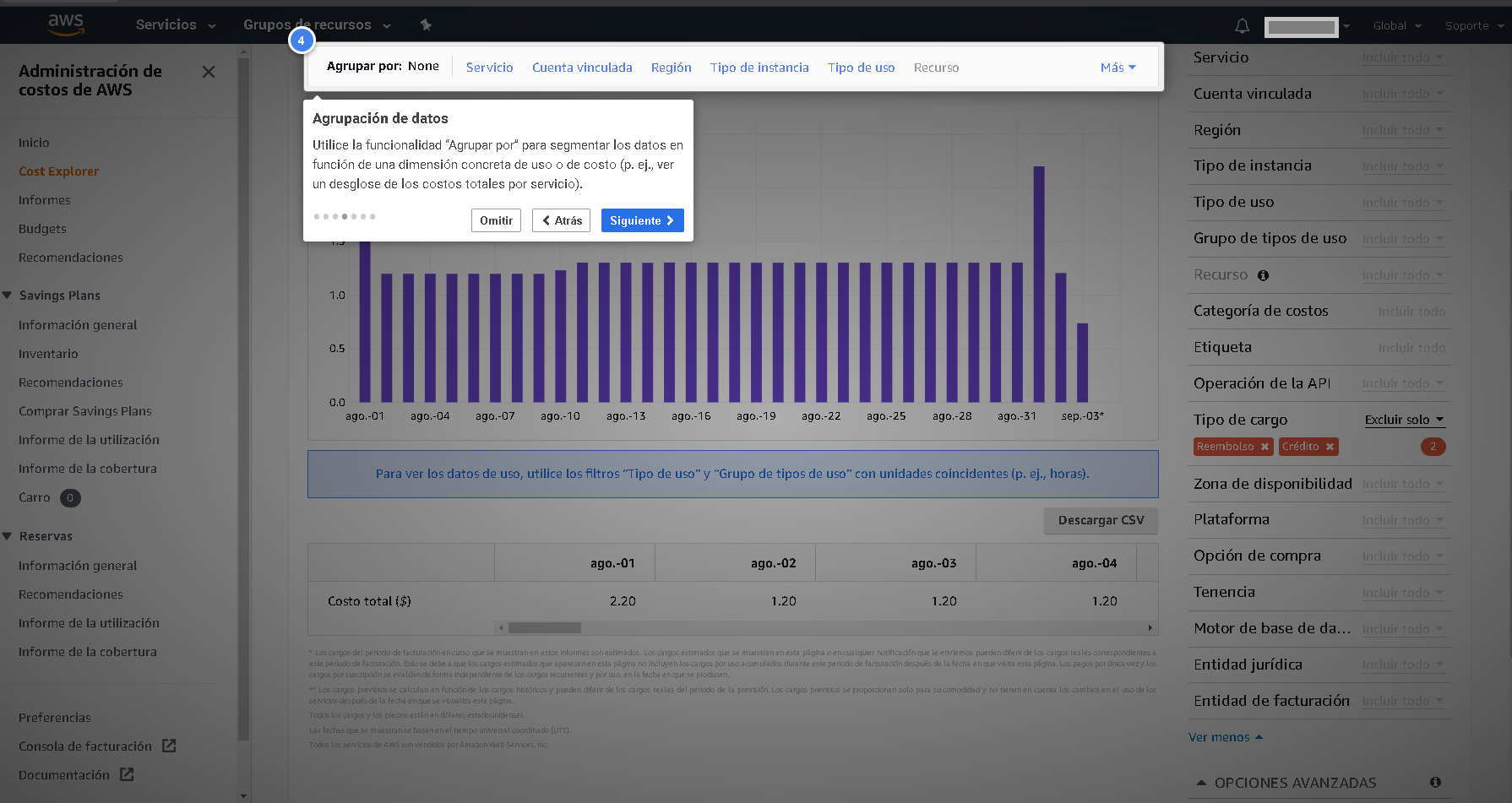
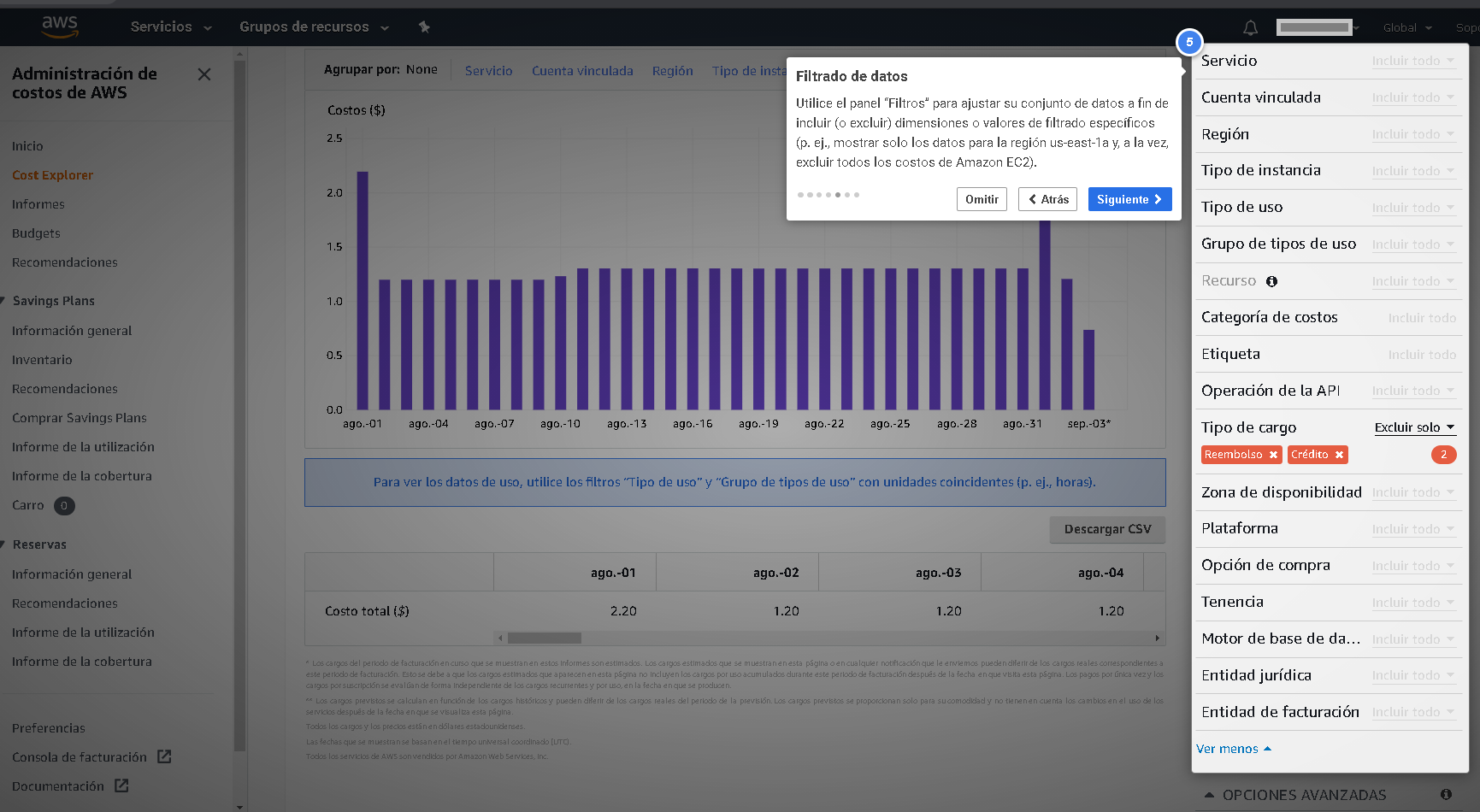
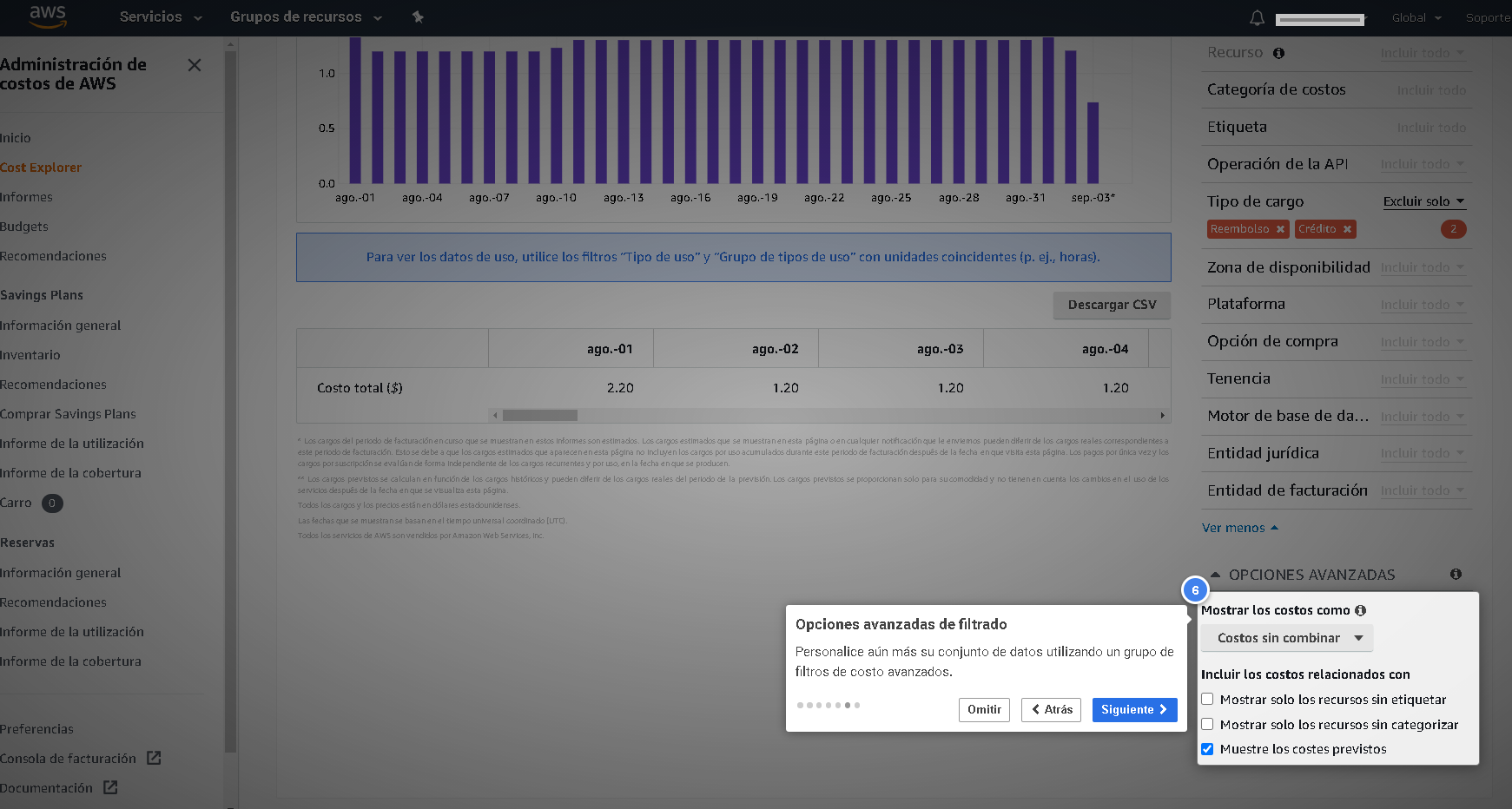
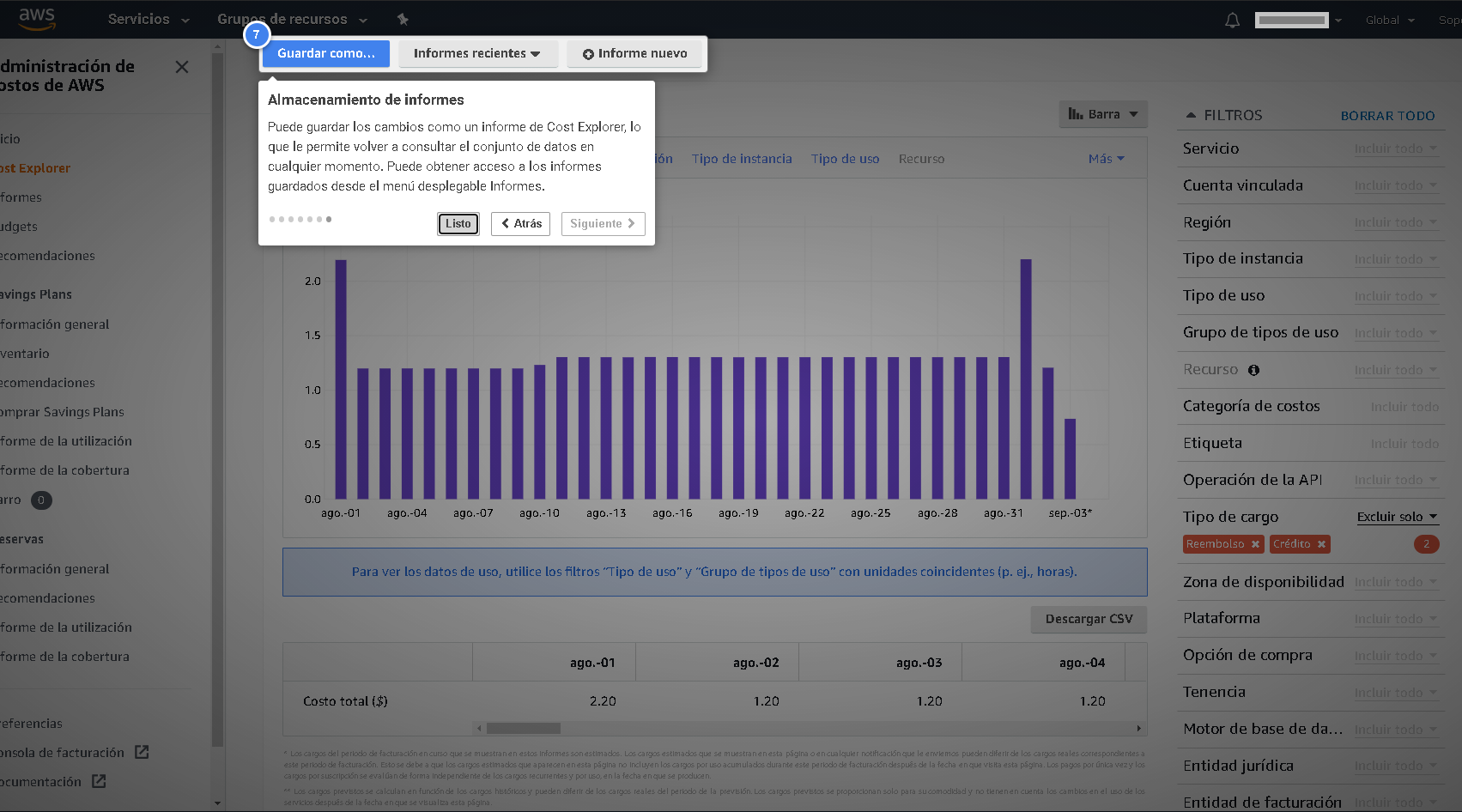
# Organización y optimización de costos en AWS

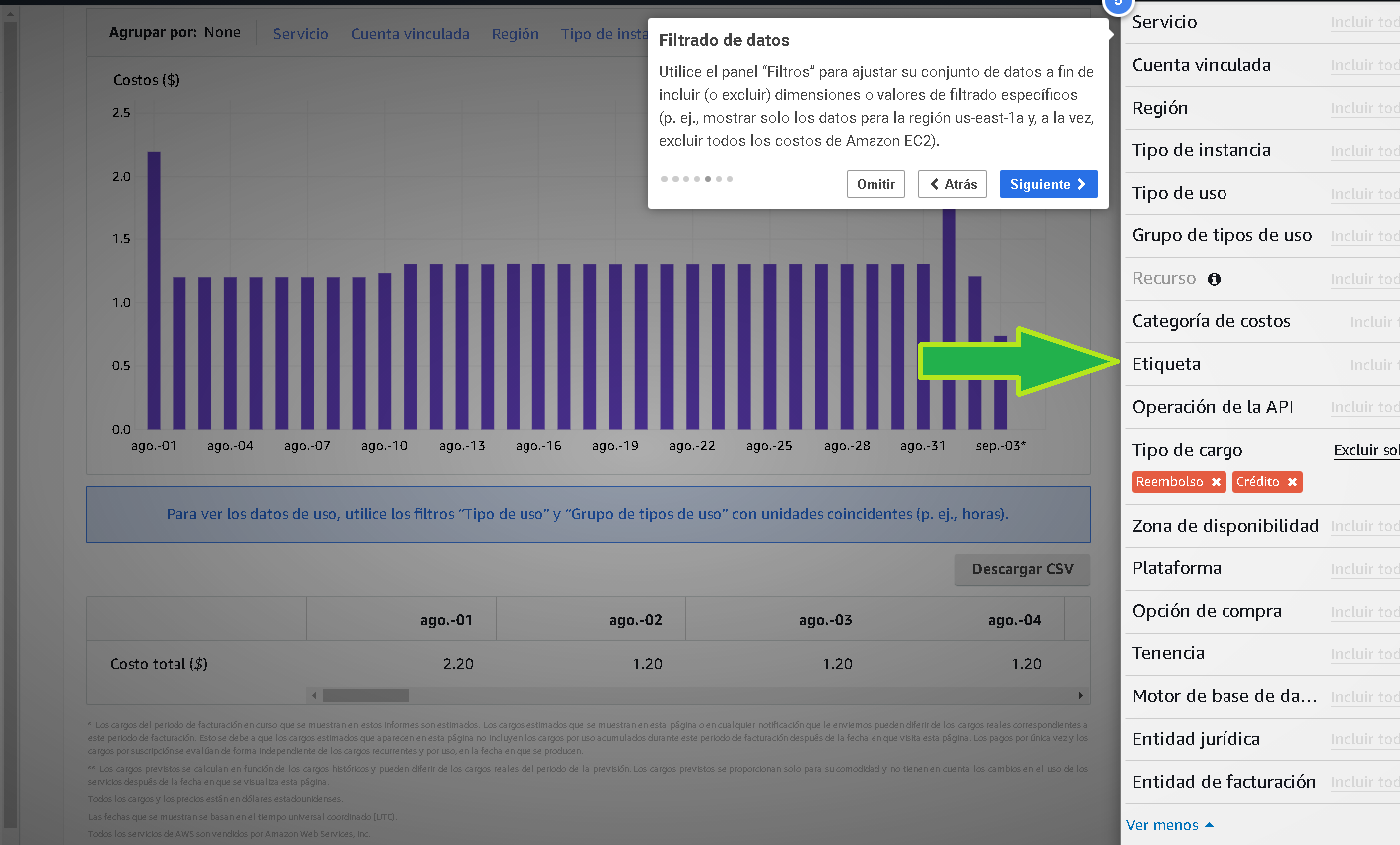
Una estrategia de ahorro económico es imprescindible a la hora de tener recursos de cómputo ya sea en la nube o en sitio. Siempre que se habla de optimización es inherente hablar de medir (lo que no se mide no se mejora). AWS ofrece un par de herramientas importantes a la hora de medir los costos que generados.

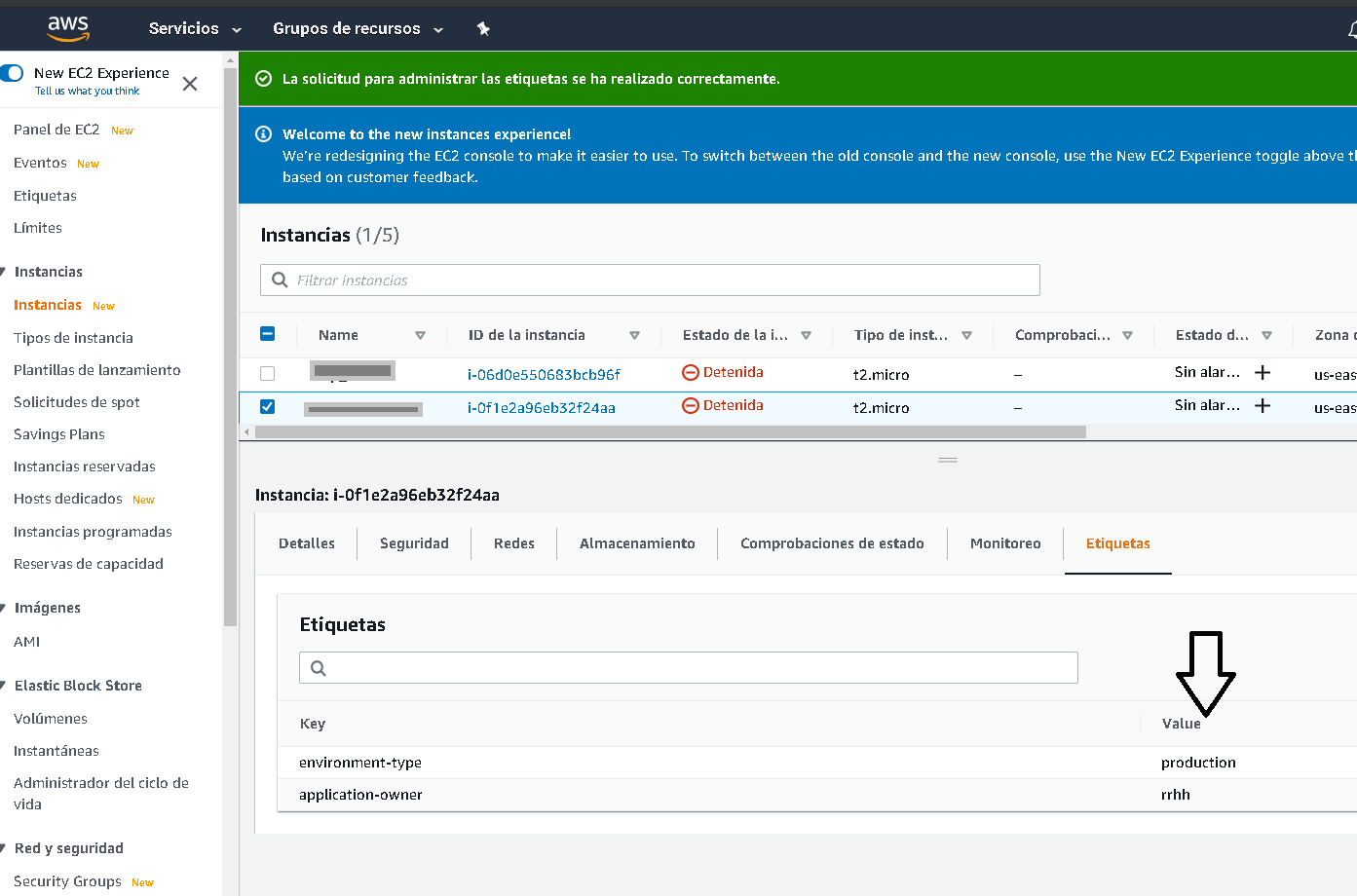
La primera es AWS Cost Explorer, la segunda no es como tal una herramienta de AWS, es una funcionalidad de AWS para poder identificar los recursos usados. La siguiente es un panel real de cost explorer. 

En el panel se puede ver la información de costos de diferentes maneras. 

EL panel de exploración de costos tiene las siguientes funcionalidades:

Y en conjunto con Tags de los recursos se vuelve una herramienta poderosa a la hora de ver los costos de los servicios en AWS usados. 

¿Qué son las etiquetas? Se pueden poner tantas etiquetas como se requieran a los recursos, con la premisa básica que sean consistentes, si en un lado se nombran los servicios como production, en otro prod y en otro producción hará difícil en análisis de costos. Las etiquetas al final son texto que se puede adherir a los recursos para identificar los recursos usados, es equivalente al mundo real pegar una etiqueta con un identificador como nombre a un servidor por mencionar un ejemplo. 

¿Qué estrategias usar para aminorar costos?

Una estrategia para aminorar costos es la reserva de instancias (RI). Puede ser hecha desde un pago inicial pago intermedio o pago final siempre y cuando haya un compromisos a pagar de 1 a 3 años de instancia, con lo que se puede ahorrar un 69% respecto a los costos bajo demanda. Instancias permitidas permitidas para ser reservadas son de tipo Amazon EC2, Amazon RDS, Amazon Redshift, Amazon ElastiCache y Amazon Elasticsearch.

Otro método para ahorro de dinero es hacer uso de un Saving Plan, los saving plans permiten ahorrar hasta un 72% de costos sobre instancias bajo demanda, en los saving plans se debe especificar un compromiso de gasto por hora de una familia de instancias por tiempos de 1 a 3 años, ej: adquirir un saving plan comprometiendo gastar 1 USD cada hora como mínimo por un periodo de 1 o 3 años.

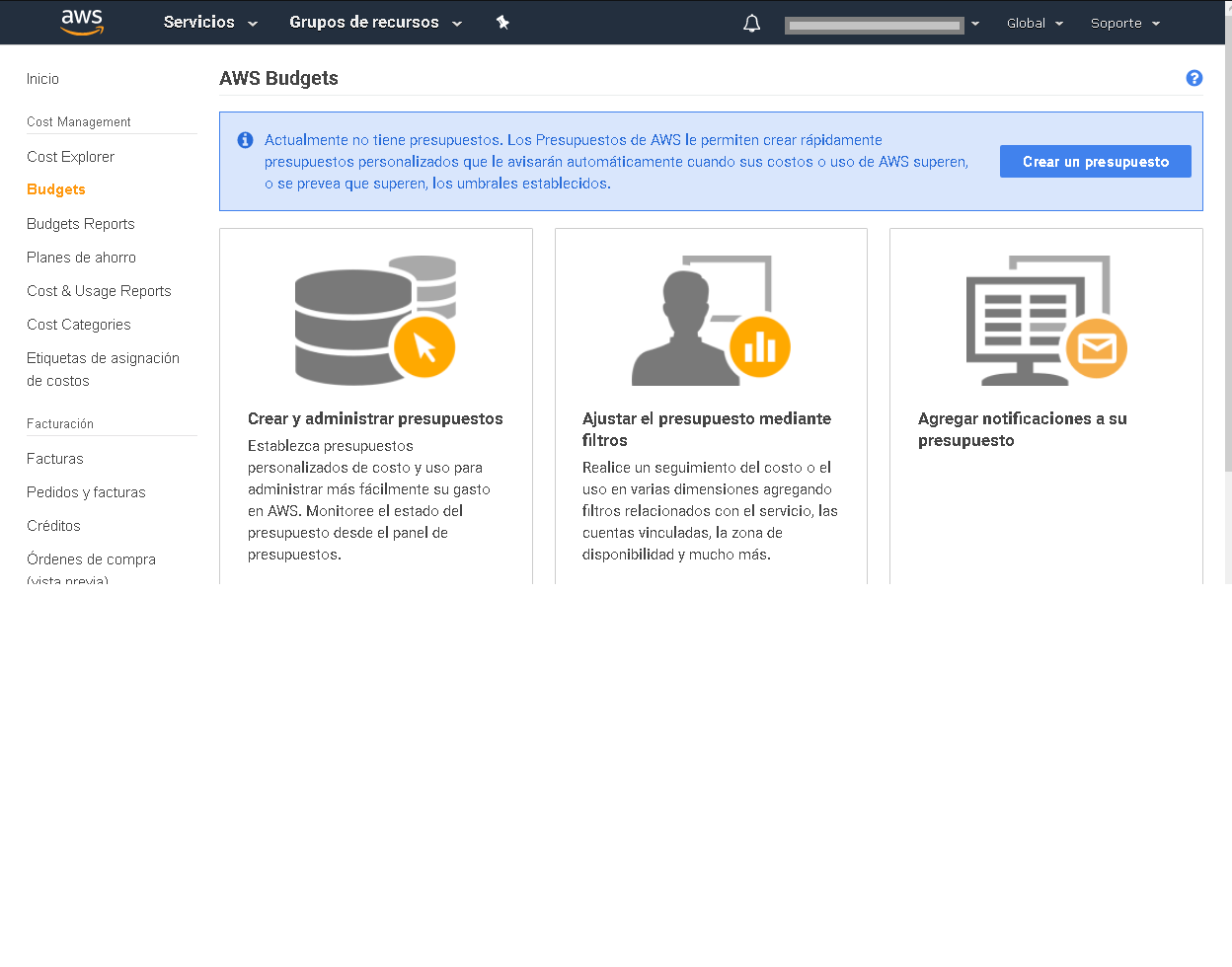
# Creación de presupuestos con AWS Budget

Los budgets o presupuestos permiten tener alarmas para el control de los gastos en los servicios de AWS evitando sorpresas en la facturación a final de mes. Es posible definir presupuestos en base a costos donde se establece una cantidad monetaria máxima que se desea gastar.

También es posible establecer alarmas de presupuesto basado en uso, se miden en unidades dependiendo el servicio, en el caso de EC2 se mide por hora de instancia encendida, con S3 en el número de requests a los buckets o en GB para la transferencia de datos, ya dependerá del servicio que se desea monitorear la unidad de medida a utilizar.

En el caso de querer monitorear las instancias reservadas se pueden generar presupuestos de reserva, con ellos se puede medir el porcentaje de utilización de la reserva de instancia se tiene consumida o que porcentaje de cobertura se tiene.

En el caso de presupuestos de saving plans se pueden medir y lanzar alarmas ante un porcentaje de utilización que se defina o bajo un porcentaje de cobertura.



budgets

Es importante mencionar que los presupuestos no [son gratis](https://aws.amazon.com/es/aws-cost-management/pricing/), a septiembre del 2020 cuestan 0.02 USD por día cada uno.

# AWS TCO Calculator, AWS Simple Monthly Calculator & Cost

Al explorar la posibilidad de una migración hacia la nube, muchas veces hay un desacuerdo sobre el costo total de propiedad de los elementos de hardware y de centros de datos existentes. AWS proporcionan una calculadora TCO flexibles para comparar el costo de funcionamiento de una puesta a punto en local (o baremetal) frente a sí mismo en Amazon Web Services. A la hora de planificar el movimiento hacia la nube muchos departamentos de la empresa están involucrados en la decisión. Operaciones e ingeniería quiere mirar y desplegar rápidamente mientras finanzas quiere gastar menos. La calculadora de AWS puede ayudar a añadir un poco de rigor numérico para el debate entre áreas.

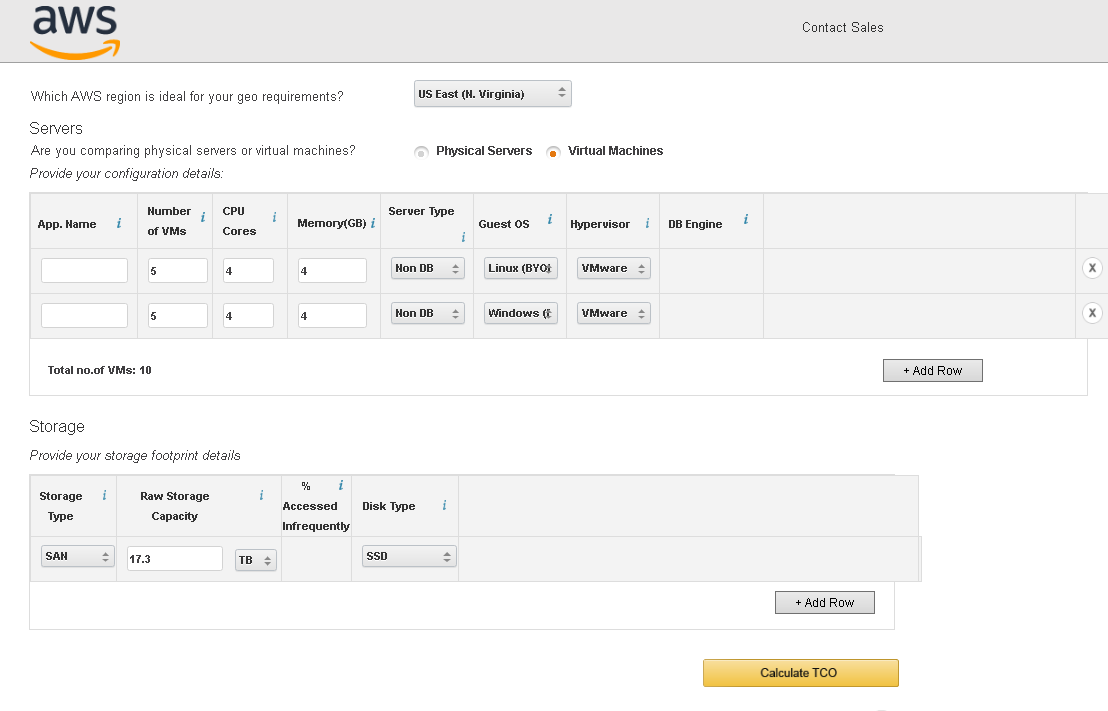
No es una tarea fácil calcular el costo de migración, algunos de los servidores no necesariamente se encuentran trabajando 24/7, o trabajando con un rendimiento de CPU del 100%, por lo que se puede estar subdimensionando o sobredimensionando.

La [calculadora](https://awstcocalculator.com/#) solamente dará un estimado de muy alto nivel, un número de referencia para partir hacia análisis mas profundos y certeros.

Para estimación de costos de uso mensuales se utilizaba [SIMPLE MONTHLY CALCULATOR](https://calculator.s3.amazonaws.com/index.html) la cual resultó en muchas veces ser una herramienta poco precisa al dejar de lado múltiples variables a calcular.

La nueva versión es [AWS Pricing Calculator](https://calculator.aws/#/createCalculator), esta nueva calculadora contempla más variables para hacer cálculos, sí, hay que conocer mas información a la hora de usarla pero se convierte al final en una herramienta mas precisa que evitará sorpresas amargas al final de mes en la factura.

Para ilustrar mejor imaginar el ejercicio: Definamos un centro de datos empresarial, con servidores on premise con 10 servidores físicos, cada unos con unas 10 máquinas virtuales cada uno en promedio con 4 GB de RAM, 50 GB de disco duro alojados en una SAN con capacidad de 20 TB en arreglo RAID 5 por lo que serán usables unos 17.3 TB solamente, 2 sockets para procesador, con procesadores Intel® Xeon® E5 2600 v4 que tienen 22 núcleos aunque cada maquina virtual usa 4 cores, la mitad con sistema operativo GNU/Linux y otras con Windows, todas ejecutandose en hypervisores de nivel 1 de VMWare,

La calculadora sería llenada con los siguiente parámetros: 

Al final el reporte proporcionado es muy completo, contiene inclusive la descripción de la metodología usada para hacer el cálculo. [Detalle](https://tcoprd.s3.amazonaws.com/2020/2020/2020/358db9d1132436096195664305.xml?AWSAccessKeyId=AKIAJIN5MU5FZ6H45LHA&Expires=1599222692&Signature=4V1sycy7K7JRzFZMXhPCeoSIgUM%3D)

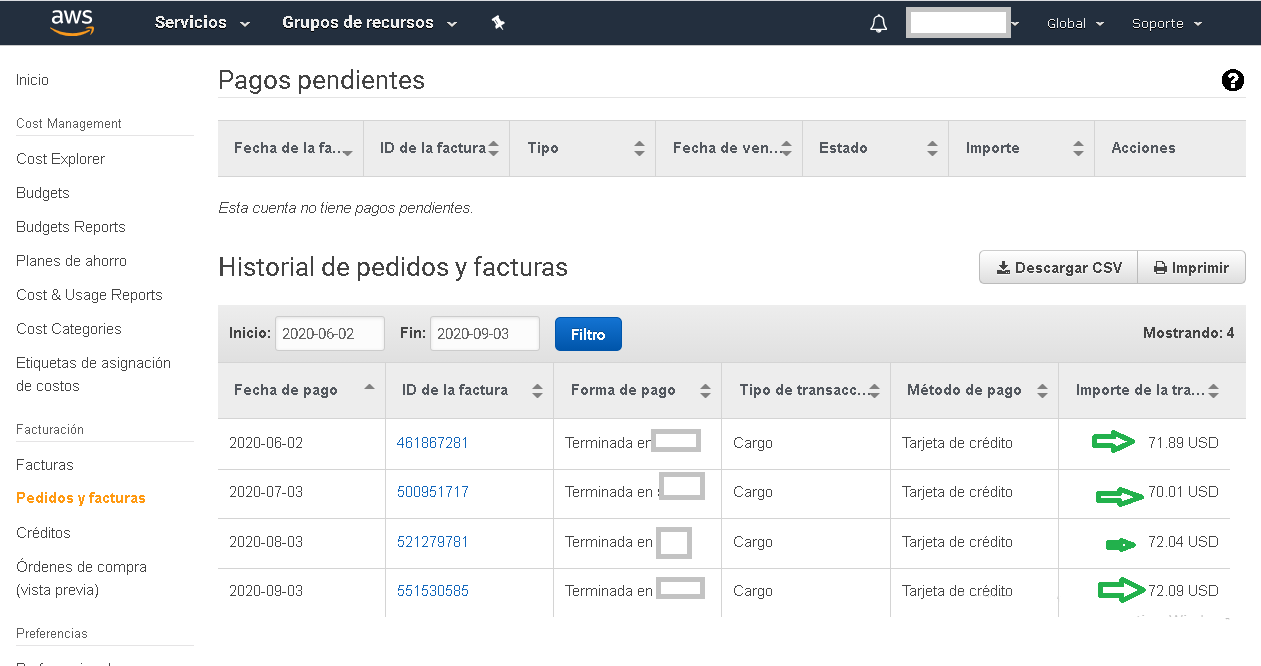
Debe tomarse este estimado con cuidado ya que los cálculos descritos en la metodología contemplan equipos que pueden diferir de precios, los sueldos de personal también difieren y los costos de energía eléctrica, por lo que no se recomienda sea esta una herramienta para tomar la última decisión.

Por otro lado ya operando servicios en AWS surgirá la duda en cuanto a costos de tener una nueva aplicación funcionando, muchas veces es para poder brindar un precio a un tercero que quiera usar un servicio ofrecido por la organización o para el departamento financiero internamente a modo de proyección. Para este caso AWS Pricing Calculator es la herramienta de estimación de costos idónea. Seguir el siguiente ejercicio: Simular una aplicación web de registro de personal de baja carga (pocas altas y bajas al día), conformada por un servidor EC2 y una base de datos de RDS, además la organización cuenta con otro servidor EC2 ejecutando un LMS(Moodle).

Ambas instancias son del tipo t2.small, tiene una transferencia hacia internet de no mas de 1 GB al mes, la base de datos es de RDS es de tipo t3.small Maria DB, con un disco de 30 GB.

AL final la [calculadora](https://calculator.aws/#/estimate?id=dc483dba879f7338a02be8bcf2c50959d53ad9f1) da un precio al mes de 78.80 USD

Esta infraestructura en producción al mes tiene el siguiente histórico de billing que es sorprendentemente similar.



calculadora-mensual