aws ACK STORM zesty

WHITEPAPER

Optimización de costes en AWS desde cero

Descubre las claves de la optimización de costes en la nube de Amazon Web Services con iniciativas detalladas, herramientas útiles, tácticas y estrategias de mapeo que te ayudarán a obtener más por tu inversión en cloud.

Contenidos

01 Introducción

02 Entender el problema del aumento de los costes de la nube de AWS

03 Cómo superar los principales retos de la optimización de costes

04 Servicios para la optimización de costes

05 Programa de precios privados en AWS

06 Modelos de precios en AWS

07 Sobre ACKstorm

Nota: este documento es una adaptación abreviada y traducida al español por ACKstorm del whitepaper de Zesty *Cost Optimization on AWS: From Zero to Hero*.

01 Introducción

La optimización de costes en la nube no es fácil; sólo el 15% de las compañías que aplican estrategias FinOps han conseguido que sea una práctica "madura y en evolución". El 75% restante todavía está en proceso de averiguar cómo fomentar y hacer evolucionar sus estrategias de optimización de costes para disfrutar de un mayor impacto en su cuenta de resultados.

¿Qué significa esto para los equipos DevOps actuales, que necesitan poder construir libremente en la nube para mantener el ritmo del cambio, pero que también necesitan ser responsables ante la espiral de costes de la nube?

En este whitepaper, elaborado por Zesty & ACKstorm, trataremos estos 3 aspectos clave:

- **Entender el problema: desde los recursos no necesarios a los no optimizados, ¿por qué se descontrolan los costes de la nube? Enumeraremos los principales problemas de las organizaciones actuales.**
- **Superar los retos: veremos consejos para reducir costes e implantar las mejores prácticas de optimización en la nube, además de herramientas y productos prácticos que pueden reducir la presión en los costes.**
- **Implantación de la automatización: cómo con tecnología más reciente utiliza la IA para automatizar el aprovisionamiento de la nube, tomar decisiones inteligentes en tiempo real y, en última instancia, ahorrar mucho dinero.**

02 Entender el problema del aumento de los costes de la nube de AWS

Con tanta gente citando el ahorro de costes como la razón principal para trasladar la infraestructura a la nube, puede resultar confuso entender por qué el gasto en cloud es un problema tan importante para las empresas de hoy en día.

La verdad es que, aunque el ahorro de costes es uno de los dos principales impulsores de la migración a la nube para el 32% de las organizaciones, una vez que se llega al cloud y se empieza a escalar según madura el crecimiento de la empresa, el rendimiento de la inversión tiende a disminuir e incluso a estancarse. El 68% de los responsables de la toma de decisiones de IT afirman que su infraestructura en la nube se ha encarecido en el último año, y la falta de mejores prácticas junto con los costes innecesarios podrían estar dejando sobre la mesa 500.000 millones de dólares sin explotar.

La creación de un equipo de FinOps, la identificación de las principales partes interesadas, la implantación de procesos de gobernanza y la utilización de una herramienta CFM contribuirán a garantizar que los costes se controlan y gestionan adecuadamente.

AWS recomienda 5 prácticas para optimizar los costes de la arquitectura en la nube. Si fallas en una o dos de estas prácticas, es fácil que lo veas reflejado en tu factura mensual. Vamos a repasarlas.

Gestión financiera en la nube: la naturaleza dinámica del cloud dificulta el seguimiento de los costes si no existen los procesos y procedimientos adecuados.
El desarrollo de un equipo de FinOps, la identificación de las principales partes interesadas, la aplicación de procesos de gobernanza y la puesta en marcha de una herramienta CFM para gestionar las instancias reservadas o el escalado automático contribuirán en gran medida a garantizar que los costes de la nube se controlan y gestionan correctamente.

Visibilidad del gasto y del uso: la flexibilidad del cloud permite a varios equipos de la organización utilizar los servicios en la nube simultáneamente, lo que dificulta la obtención de una visibilidad holística del uso. La escasa integración entre los sistemas en la nube y la falta de alineación entre las herramientas y tecnologías en cloud que los equipos despliegan a menudo da lugar a la duplicación de datos, esfuerzos y gastos en varios equipos. El etiquetado (“tagging”) de recursos, la configuración de herramientas de supervisión y el uso de herramientas como AWS Cost Explorer pueden ayudar a obtener una mayor visibilidad del gasto.

Recursos con mayor rentabilidad: AWS ofrece una variedad de modelos de precios con mayor rentabilidad para ayudar a sus clientes a obtener más por su dinero. Las instancias reservadas ofrecen hasta un 70% de descuento en el uso de EC2 bajo demanda, los planes de ahorro ofrecen un descuento menor, pero proporcionan una mayor flexibilidad si necesita realizar cambios en los compromisos de EC2, y las instancias "spot" pueden proporcionar hasta un 90% de descuento, pero sólo funcionan para el uso a corto plazo en circunstancias limitadas. También es posible que quieras considerar además, opciones para el almacenamiento de redundancia reducida (RRS). Examinaremos todos estos modelos con más detalle más adelante en esta guía.

AWS ofrece una variedad de modelos de precios rentables para ayudar a sus clientes a obtener más por su dinero.

Gestión de la oferta y la demanda de recursos: como los ingenieros están atados a un desarrollo rápido, a menudo sobre-aprovisionan en recursos, para no quedar desamparados si necesitan escalar rápida o inesperadamente para satisfacer la demanda. Este despilfarro de recursos en la nube tiene un grave impacto en la cuenta de resultados, y es un problema común. De hecho, el 66% de las organizaciones admitieron haber aprovisionado en exceso el año pasado. Por otro lado, el aprovisionamiento insuficiente también es un problema, ya que los desarrolladores se quedan sin los recursos que necesitan, a menudo debido a las estrategias de optimización que intentan hacer frente a la espiral de coste.

El 66% de las organizaciones admite haber aprovisionado en exceso el año pasado.

Optimización continua: a medida que cambian las necesidades, si no se eliminan los recursos no necesarios o no utilizados, si no se tiene en cuenta los residuos de la nube y no se implementan nuevas funciones y tipos de recursos puede salir caro. La mayoría de las organizaciones admiten tener recursos no necesarios o poco o nada utilizados en su sistema, que pueden ser cualquier cosa, desde recursos que funcionan a tiempo completo aunque no se utilicen durante más de unas horas, o instancias olvidadas que se utilizaron para entornos de desarrollo o pruebas que no se eliminaron después de su uso.

Normalmente los recursos inactivos o infrautilizados son un fallo en la gestión de los recursos en la nube. Muchas veces esto ocurre por falta de formación en torno a la gestión de un entorno cloud, por ejemplo cuando no se programan la destrucción o la parada de instancias y se dejan funcionar las 24 horas del día. Otras veces puede ser intencionado, porque los equipos de DevOps han sobre-aprovisionado para no quedarse sin los recursos que necesitan.

03 Cómo superar los principales retos de la optimización de costes

La preocupación por la optimización de costes es común en toda la comunidad cloud, por lo que ya existen soluciones que pueden traducirse en importantes ahorros.

La primera recomendación es buscar cómo adaptar la arquitectura para tener en cuenta los costes siempre que sea posible. Se trata de un “entretenido” ejercicio de equilibrio entre contener y medir el riesgo, al mismo tiempo que se asegura el avance y se sigue el ritmo del negocio.

Por ejemplo, es posible que quieras modificar las configuraciones de red o regiones de AWS para obtener un mejor precio en la transferencia de datos. Por ejemplo EC2 cuesta un 11% menos en la región de Estocolmo que en la de Londres, por lo que podría ser un movimiento inteligente desde el punto de vista del ahorro de costes. Pero espera. ¿Este cambio afectará a la disponibilidad, regulación y el rendimiento de tus aplicaciones?

Es necesario tener esa conversación antes de pulsar el botón de lanzar. La clave está en encontrar la manera de hacer que el proceso sea menos manual y más automatizado.

Por otro lado, la propiedad y responsabilidad son absolutamente críticas cuando se considera la gestión financiera en la nube, razón por lo que establecer un Centro de Excelencia Cloud (CCoE) es una muy buena práctica.

Según Gartner, "para garantizar el éxito de la adopción de la nube, las organizaciones deben contar con las competencias y la estructura adecuadas. La mejor manera de conseguirlo es crear un centro de excelencia centralizado en la nube (CCoE). Un CCoE es una función de gobierno centralizada para la organización y desempeña un papel consultivo para IT, la unidad de negocio y los consumidores de servicios en la nube de la empresa.
Un CCoEes clave para impulsar la transformación de TI en la nube".

Un CCoE establecerá quién es el responsable de la nube dentro de la empresa y garantizará que haya una propiedad clara sobre los recursos de la nube. También puede incentivar a los ingenieros para que tomen el control de sus procesos, aplicando estrategias y tecnologías rentables como parte de la construcción desde el primer día y desactivando los recursos que no se utilicen.

Ofrecer formación para incentivar a los ingenieros a que se hagan cargo de sus recursos es otra pieza fundamental del rompecabezas.

Un CCoE establecerá quién es el responsable de la nube dentro de la empresa y garantizará que existe una propiedad clara sobre los recursos cloud.

Empodéralos para que desempeñen un papel activo en el proceso de FinOps y recordarles que FinOps no significa reducir su presupuesto, sino reasignarlo allí donde aporte mayor valor estratégico. Esto significa que muchos de ellos podrían tener más presupuesto, si pueden demostrar su necesidad, una vez tengan los recursos bien optimizados.

04 Servicios para la optimización de costes

El equipo de ingeniería no puede abordar el control de costes por sí sólo. AWS ofrece de forma nativa herramientas de “Business intelligence”, el escalado automático, la detección de anomalías y monitorización. Éstas herramientas ofrecen información sobre el uso de la nube, como la detección de recursos no utilizados o innecesarios, volúmenes de EBS sin asignar o las instancias de EC2 que acumulan el mayor gasto.

La automatización es un buen aliado para la optimización de costes.

AWS Cost Explorer es un buen comienzo, ya que ayuda a visualizar y administrar el gasto de AWS a lo largo del tiempo con informes predeterminados tanto para servicios como para recursos. La automatización es un buen aliado para la optimización de costes, y puede evitar jugar a las adivinanzas cuando se afronta la compra de instancias y compromisos, así como ayudar a autoescalar los volúmenes de EBS, o dimensionar correctamente los recursos de computación.

Tomemos como ejemplo AWS Auto Scaling, que nos permite establecer políticas inteligentes para lanzar o destruir instancias a medida que cambia la demanda. También se puede configurar para que utilice diferentes tipos de instancias. Para mantener el mínimo en instancias reservadas o con “savings” y escalar usando instancias del tipo “spot”, y en caso que no estén disponibles utilizar “on-demand” (También es aplicable a clusters de EKS). Estos factores combinados ayudan en la optimización de los ahorros de costes.

05 Programa de precios privados en AWS

Existe un programa de descuentos para empresas de AWS (AWS Enterprise Discount Program), ahora llamado PPTS (Private Pricing Term Sheet). Si su organización es un cliente empresarial de AWS suscrito a Enterprise Support y puede comprometerse a un elevado consumo en cloud, especialmente durante un largo período de tiempo, AWS ofrece un descuento a cambio de un compromiso de consumo anual prepagado.

Hay un compromiso mínimo de 500.000 dólares al año, y aunque los descuentos se negocian con AWS caso por caso, en términos generales, es más interesante cuanto más largo sea el plazo de compromiso y mayor sea el gasto.

La clave está en no comprometerse demasiado con el uso de la nube y gastar más de lo necesario y, al mismo tiempo, asegurarse de no comprometerse menos de lo necesario y no aprovechar todos los descuentos. Es otro gran ejercicio de equilibrio.

06 Modelos de precios en AWS

Si estás utilizando muchas instancias bajo demanda (“On-Demand”). Estas son algunas de las principales formas de reducir costes a través de otros modelos predefinidos de AWS:

- Instancias reservadas (Reserved Instances - RI)
- Planes de ahorro (Saving Plans - SP)
- Instancias tipo “Spot”

Instancias reservadas (Reserved Instances - RI): se trata de un sencillo descuento en la facturación de los costes de uso de EC2. Es necesario comprometerse a un plazo de contrato de uno o tres años y, a continuación, se le factura por cada hora, independientemente de si utiliza o no la instancia.

El ahorro estándar en RI es de una media del 40% para un compromiso de un año, y aumenta al 60% para un plazo de tres años. Por lo tanto, es importante considerar lo que va a necesitar con antelación. Las RIs no se pueden cancelar y, por lo general, no se pueden modificar (solo ciertos atributos). Aunque pueden venderse en el mercado de RIs de EC2 si ya no las necesita, normalmente el precio de venta es muy inferior al del mercado.

Planes de ahorro (Saving Plans - SP): existen dos tipos diferentes de planes de ahorro, uno para computación y otro para Instancias EC2.

Los planes de computación son mucho más flexibles y pueden utilizarse para EC2, Lambda y Fargate, en cualquier familia, tamaño, región o sistema operativo. Los planes de ahorro de instancias EC2, requieren un compromiso con un único tipo de familia de instancias en una región. En pocas palabras, comprometerse con un tipo de instancia concreto en una región específica de AWS durante un año o tres años para obtener un descuento en el gasto bajo demanda de hasta el 72% en EC2 y de hasta el 66% en el plan de computación.

IInstancias tipo “Spot”: las instancias puntuales son la forma más económica de comprar capacidad EC2, pero los casos de uso son limitados para la mayoría de las empresas, ya que AWS puede solicitar la devolución de la capacidad con un preaviso de dos minutos. Si tiene aplicaciones “stateless” y, tolerantes a fallos o flexibles, podría disfrutar de un descuento del 90% sobre los precios habituales. Las cargas de trabajo de entornos de prueba o desarrollo son una buena opción, también los modelos de entreno en big data, donde una hibernación con poca antelación no es un problema. Recuerde que las cargas de trabajo deberán configurarse para poder aprovecharlas, por lo que tendrá que tener en cuenta el esfuerzo adicional de ingeniería.

07 Sobre ACKstorm

ACKstorm es un partner especializado en soluciones cloud nativas, centrado en diseñar, construir, gestionar y modernizar infraestructuras en la nube de AWS y Google Cloud. Colaboramos estrechamente con nuestros clientes, esforzándonos por mejorar el valor comercial a través de tecnología de nube de última generación.

Priorizamos la construcción de relaciones basadas en la confianza mutua, haciendo hincapié en la transparencia y honestidad. Nuestro compromiso incluye proporcionar soluciones con enfoques sin bloqueo de socios y sin preferencia de proveedor.

Descubre más en www.ackstorm.com

Si tienes cualquier duda o quieres conocer más detalles de cómo te podemos ayudar:

info@ackstorm.com

+34 93 393 63 88 | +34 91 005 41 66

ACK STORM