**Resumen:** los consumidores pueden obtener importantes ahorros de costos comprando irecursos de mercados contables computacionales como las instancias puntuales de Amazon Elastic Compute Cloud (EC2). Pueden aprovechar las diferencias de precio entre ranuras de tiempo, regiones y tipos de instancia para minimizar el costo total de ejecutar sus aplicaciones en la nube. Sin embargo, los mercados Spot son inherentemente volátiles y dinámicos, como consecuencia de esto, los precios spot cambian continuamente. Como tal, los postores potenciales pueden beneficiarse de una visión inteligente de la dinámica del mercado Spot que puede ayudarlos a tomar decisiones de oferta más informadas. Para permitir esto, proponemos un enfoque estadístico descriptivo para el análisis de los mercados de Amazon EC2 Spot para detectar patrones de fijación de precios típicos, incluida la presencia de componentes marinos, extremos y tendencias. Utilizamos **tres medidas estadísticas: el coeficiente de Gini, el índice de Theil y el promedio móvil ponderado exponencial**. También diseñamos un modelo para estimar ofertas mínimas de modo que las instancias de Spot se ejecuten durante duraciones específicas con una probabilidad mayor que un valor establecido en función de diferentes periodos de retrospectiva. Los resultados experimentales muestran que nuestra estimación arroja como promedio una estrategia de oferta que puede asegurar de forma confiable una instancia al menos el 80% del tiempo con una garantía mínima objetivo entre 50% y 95%.

1. INTRODUCCIÓN

Una característica clave de la computación en la nube es el acceso percibido ilimitado a demanda a los recursos informáticos. Los proveedores generalmente logran esto mediante el aprovisionamiento en exceso de sus centros de datos para satisfacer la demanda máxima; en consecuencia, tienen que luchar frecuentemente con la capacidad no utilizada. **Dado que la infraestructura de la nube es un producto perecedero, los proveedores requieren técnicas apropiadas de gestión del rendimiento para minimizar el capital inactivo y lograr una mejor utilización**. Como pionera de la computación en la nube, **Amazon Elastic Compute Cloud (EC2)** **utiliza tres opciones de compra** diferentes para mejorar la administración del rendimiento: **instancias reservadas, instancias bajo demanda e instancias puntuales.**

* Con las instancias reservadas, los consumidores incurren en un costo fijo relativamente alto, pero tienen acceso garantizado a pedido, y las instancias no son anuladas una vez aprovisionadas.
* Con las instancias On-Demand, los consumidores incurren en un alto costo fijo y no tienen acceso garantizado a pedido, pero las instancias no son anuladas una vez aprovisionadas.

Por lo tanto, Amazon ofrece garantías de disponibilidad mínima sostenida con estas dos opciones de compra.

Con Spot Instances, Amazon permite a los usuarios pujar por la capacidad no utilizada, pero no ofrece garantías sobre la disponibilidad mínima sustenida. **Los usuarios envían solicitudes de instancia puntual especificando el precio máximo por hora que están dispuestos a pagar por los recursos requeridos**. El precio Spot se determina ejecutando una subasta de mercado sellado, multiunidad y precio uniforme. **A los usuarios se les asignan instancias puntuales si su oferta es más alta que el precio Spot**. Sin embargo, **Amazon termina las instancias puntuales siempre y cuando el precio spot exceda el precio máximo de oferta**. También puede revocar instancias puntuales si el tamaño del Spot Pool disminuye, es decir, el suministro de instancias puntuales disponibles disminuye debido a una mayor demanda de instancias bajo demanda y reservadas, y el precio máximo de oferta es igual al precio spot. Al igual que en otros mercados de productos básicos, Amazon EC2 Spot Market también experimenta períodos de estabilidad y volatilidad. Esto, junto con la opacidad de los mercados de EC2, complica el proceso de licitación de las instancias puntuales. **Los usuarios pueden elegir entre hasta 59 tipos de instancias diferentes que se pueden ofrecer en 31 zonas de disponibilidad y ejecutar 6 sistemas de operación diferentes**. Por lo tanto, en principio, un consumidor normal de EC2 puede elegir **entre 10,974 mercados al contado únicos**. El número real de mercados al contado es ligeramente inferior, ya que Amazon no ofrece todos los tipos de instancias que ejecutan todos los sistemas operativos en todas las zonas de disponibilidad. Con el fin de ayudar a los usuarios con sus estrategias de oferta, Amazon ofrece hasta 90 días de datos históricos de precios directos para cada mercado mediante una Interfaz de programación de aplicaciones (API) accesible a través de la web y a través de la Consola de administración EC2. También proporciona una aplicación Spot Bid Advisor que ayuda a los usuarios a determinar los precios de oferta en los diferentes mercados según sus necesidades.

**Sin embargo**, *analizar el historial de precios de Spot para una cantidad tan grande de mercados, entender la dinámica de precios dentro y entre ellos, y determinar el mejor tipo de instancia, zona de disponibilidad, intervalo de tiempo y precio de oferta* **no es sencillo.** **Los postores potenciales necesitan ayuda** para dar sentido a la creciente cantidad de historial de precios Spot que está disponible para ellos. **Pueden beneficiarse** de un servicio que puede analizar e interpretar el historial de precios de Spot y **(a) cuantificar la volatilidad y variabilidad de cada mercado Spot, (b) identificar patrones de precios y tendencias dentro y entre mercados**, y **(c) calificar el riesgo asociado con pujas Spot**. En este artículo, **utilizamos un enfoque de estadística estático basado en tres medidas estadísticas** diferentes para lo cuantitativo. En particular, utilizamos dos índices econométricos de desigualdad, a saber, el **coeficiente de Gini y el índice de Theil** para comprender los patrones de precios del mercado Spot EC2. La aplicación de estos dos índices a los datos históricos del precio Spot nos ayuda a **detectar dos características**: **(i) ocurrencias de extremos del precio Spot, y (ii) la presencia de un componente estacional** (es decir, variaciones en los precios Spot que se deben a la hora del día) en un mercado de instancia de Spot determinado.

Sin embargo, mientras que el **coeficiente de Gini y el índice de Theil** *pueden informarnos sobre la existencia de tales patrones* (únicamente a través de un único número cada uno), *no pueden revelar la ubicación y la forma de estos patrones*. **Para comprender completamente la naturaleza de los patrones de precio spot,** utilizamos una **tercera medida**: el **gráfico de promedio móvil ponderado exponencial**. **Este gráfico** sirve como un **método de detección de cambios** que puede revelar las **tendencias del precio spot en el tiempo**, incluida la **periodicidad y la estacionalidad** indicadas por el índice Theil y los picos extremos temporarios indicados por el coeficiente Gini. Colectivamente, **estas tres medidas** dan una apreciación de los patrones de **fijación de precios en los mercados EC2 Spot**.

**También proponemos una fórmula** **para la estimación del precio oculto** basada en diferentes criterios que incluyen la duración mínima de disponibilidad sostenida, el período de retrospectiva de los datos históricos y la probabilidad de éxito de la oferta.

Realizamos experimentos en 270 mercados Spot en la región de Asia Pacífico (Sydney) para evaluar la efectividad de la fórmula propuesta. Los resultados muestran que la estimación del precio oculto es una función implícita de la presencia de componentes estacionales (capturados por el índice de Theil) y la distribución de picos extremos (capturados por el coeficiente de Gini). **Este artículo está organizado de la siguiente forma**: Presentamos un resumen del **coeficiente de Gini, el índice de Theil y el promedio móvil ponderado exponencial** en la **Sección II**. En la **Sección III**, describimos nuestro enfoque propuesto para **analizar el historial de precios Spot** utilizando estas tres medidas y presentamos un esquema para clasificar los mercados EC2 Spot basados ​​en los índices Gini y Theil. En la **Sección IV**, presentamos nuestra **fórmula de estimación del precio de oferta** para una oferta informada basada en datos históricos y la evaluamos bajo diferentes condiciones de prueba**. Presentamos el trabajo relacionado en la Sección V** y **concluimos el trabajo en la Sección VI.**

1. ESTADÍSTICAS DESCRIPTIVAS UTILIZANDO ÍNDICES DE DESIGUALDAD Y PROMEDIOS EN MOVIMIENTO

**Las estadísticas descriptivas** se refieren al análisis de distribuciones de datos que ayuda a describirlas o resumirlas de una manera significativa. Se pueden **dividir en dos categorías** amplias basadas en las diferentes características de distribución en las que se centran: las **medidas de tendencia central y las medidas de dispersión**. Si bien el uso de estadísticas de tendencia central es el enfoque más popular para el análisis univariado, que describe una distribución en términos de su valor de datos más frecuente, típico o promedio, los índices de desigualdad ofrecen una alternativa prometedora. Constituyen un método econométrico que se centra en las diferencias operativas en la distribución más que en el valor esperado de los datos. **Utilizamos ambos tipos de medidas en nuestro enfoque** **para analizar el historial de precios de Amazon EC2 Spot**. Hacemos uso de dos índices de desigualdad prominentes, a saber, el coeficiente de Gini y el índice de Theil y el gráfico de promedios móviles ponderados exponencialmente.

**A. Coeficiente de Gini**

**El coeficiente de Gini es una medida estadística de la desigualdad entre los valores de una distribución de frecuencia, como el ingreso o la riqueza**. Se encuentra dentro del **intervalo de [0, 1]** donde un coeficiente de Gini de **cero implica igualdad absoluta**, es decir, **todos los valores** en la distribución **son iguales**, y un coeficiente de Gini de **uno** (o 100%) implica la **máxima desigualdad entre los valores**. Un **beneficio** clave de usar el coeficiente de Gini es que es **independiente de tamaño y escala**. Sin embargo, **no es descomponible**, es decir, **es una medida a nivel macro que puede resumir la desigualdad a nivel de la población, pero no puede capturar la desigualdad en un nivel micro entre individuos y grupos de individuos**.

Para una población discreta con valores no negativos xi, 1 <i <n, el coeficiente de Gini es la mitad de la diferencia media relativa de cada par (xi, xi), 1 <i, j <n, en la población.

**El coeficiente de Gini** se puede calcular como una **medida resumida de los datos del precio Spot para cada mercado Amazon EC2 Spot**. Se puede usar para **caracterizar la desigualdad general en los precios Spot**.

* Un **Gini bajo** (entre 0 y 0,2) indica que el mercado Spot es homogéneo, es decir, los precios Spot siguen aproximadamente una distribución normal con una desviación muy pequeña de la media.
* Un **alto Gini** (entre 0.6 y 1.0) indica que el mercado Spot tiene picos infrecuentes pero muy altos del precio Spot promedio.
* Un **Gini de rango medio** (entre 0.2 y 0.6) indica que hay varios picos altos (pero no extremos) del precio Spot promedio.

Si bien el puntaje de **Gini proporciona una indicación de la ausencia o presencia de picos (muy) altos en los precios Spot**, **no proporciona información sobre cuándo ocurren**.

**B. Índice de Theil**

**El índice de Theil es una medida de desigualdad que se deriva de la medida de la entropía de la información** de Shannon, donde la entropía es una medida de la aleatoriedad en un conjunto dado de información. **El índice de Theil cae dentro del intervalo [0, log n]** donde un **valor alto indica una distribución desigual**. La **diferencia principal entre el índice de Theil y el coeficiente de Gini** es que mientras el **índice de Theil es descomponible, captura la desigualdad a un nivel micro entre individuos y grupos de individuos**. Permite la exploración de la estructura de la desigualdad, es decir, la desagregación de la desigualdad total a lo largo de diferentes factores. El índice Theil es un promedio ponderado de la desigualdad dentro de los subgrupos y la igualdad entre esos subgrupos, es decir, la suma de un componente entre grupos y los componentes dentro del grupo. El componente dentro del grupo captura la desigualdad debido a la variabilidad de los valores dentro del grupo. De manera similar, el componente entre grupos captura la desigualdad debido a la variabilidad de los valores en diferentes grupos.

**El índice de Theil se puede usar para verificar si un mercado Spot determinado tiene patrones recurrentes en las variaciones del precio Spot.** Los **precios al contado** se pueden **agrupar según diferentes criterios**, como la **hora del día, la ubicación física o la familia de instancias**. Para el propósito de este estudio, agrupamos los precios al contado en 24 bins utilizando **el criterio de la hora del día**. El componente dentro del **grupo del índice Theil IW (H) captura la desigualdad en los precios spot dentro de cada hora del día**, y el componente entre grupos del **índice Thiel IB (H) captura la desigualdad debido a la variabilidad en el precios spot en diferentes horas del día**. La proporción de **grupos "intermedios" RB (H)** permite el **reconocimiento de las diferencias entre los grupos**, a fin de **identificar** la presencia de un **componente estacional en el mercado Spot.**

**C. Promedio móvil ponderado exponencial**

La **Tabla de control del promedio móvil ponderado exponencial** (EWMA) **es un método para la detección de cambios que puede detectar y capturar cambios en una secuencia de observaciones aleatorias**. **Cada punto en EWMA** representa el **promedio ponderado de la corriente y un cierto número de valores observados previamente**, **con el mayor peso dado a las observaciones recientes**. Se utiliza un **parámetro A de suavizado** (también denominado factor de olvido) **para garantizar que se otorgue un peso exponencialmente menor a las observaciones que se eliminan en el tiempo**. **Cuando se aplica a la historia precio EC2** para comer, EWMA **permite la caracterización adicional de los patrones de fijación de precios** detectados por la aplicación del análisis de la desigualdad econométrico utilizando el coeficiente de Gini y la relación de grupo RB "entre" (II).

El promedio móvil ponderado exponencial se define por (mirar la fórmula)

donde A E [0, 1] es el factor de olvido que es una vari-poder aproximar la media de la xi observaciones al azar, 1 <i <n, pero la colocación de más peso a las observaciones más recientes. es el factor de ponderación aplicado a la muestra de datos actual xi, dada por (mirar la fórmula).

1. ENFOQUE DE ANALÍTICA DESCRIPTIVO

**A. Recolección de datos, consolidación y preparación**

Actualmente, **Amazon ofrece hasta 90 días de historial de precios Spot**. Si los usuarios de EC2 desean utilizar más de 90 días de datos para el análisis, tienen que crear su propio repositorio y almacenar la información histórica, que luego se puede recuperar más tarde para el análisis fuera de línea. Para habilitar esto, hemos incorporado un componente que se puede usar para la recopilación y consolidación continua del historial de precios de EC2 Spot. **Usamos la API de precio Spot de Amazon EC2** para r**ecopilar** **el historial de precios Spot** que luego se almacena en formato de series de tiempo utilizando la **base de datos InfluxDB3** de fuente abierta. **Cada registro** captura el **thnestamp** (es decir, la fecha y hora en que se produjo el cambio de precio) y el **precio Spot correspondiente** (es decir, el precio por hora establecido por Amazon EC2 en función de la oferta y la demanda de instancias puntuales). Cada **combinación** única de availabili-tyZone, instanceType y productDescription representa un **mercado EC2 Spot**. Nuestro repositorio local posee actualmente el historial de precios Spot a partir del 18 de junio de 2016. Sin embargo, **en este documento**, limitamos nuestro **análisis** a la región de **Sydney** y analizamos los **datos de precios Spot recopilados entre el 2 de julio de 2016 y el 19 de octubre de 2016**, la fecha de inicio se debe al tiempo de entrega necesario para obtener un precio Spot inicial para todas las instancias EC2 analizadas.

Los **datos históricos se preprocesan** de la siguiente manera **antes** de que se pueda realizar el **análisis estadístico** en él:

- **El historial de precios Spot** recuperado de **InfluxDB** se convierte del **formato json al formato csv** utilizando un analizador personalizado.

- **Los datos se dividen en diferentes contenedores según el criterio de la hora del día y calculan el precio máximo por hora por cada hora**. La interconexión es un requisito previo para la **aplicación del índice Theil** y para **determinar si los datos del precio Spot contienen un componente estacional**. Aunque los precios Spot pueden cambiar varias veces en una hora, solo retenemos el precio Spot máximo observado por cada hora.

- **La naturaleza del mercado Spot EC2** es tal que un precio Spot puede **permanecer constante por largos períodos de tiempo**. Dado que **los precios Spot se registran solo cuando cambian**, los datos históricos pueden no contener información del precio Spot para todas las horas del día. Para tener en cuenta este comportamiento, sintetizamos los datos faltantes. **Sustituimos** todos los **datos faltantes** secuencialmente con el último precio máximo registrado para que no haya **N / A** (quitar los nans) en los datos.

Al final de la fase de preparación de datos, el conjunto de datos consta de 1,77 millones de registros para la región de Sydney, con un total de 3 sistemas operativos, 34 tipos de instancias y 270 mercados de instancias Spot distintos (90 por sistema operativo).

**B. Análisis econométrico de datos de precios spot**

Los **índices econométricos** de desigualdad, como el **coeficiente de Gini y el índice de Theil**, han sido desarrollados por investigadores económicos para m**edir la desigualdad en la distribución del ingreso dentro de una sociedad**. Más precisamente, en los estudios econométricos, estos índices **capturan las disparidades entre la cantidad de recursos** (es decir, el ingreso en porcentaje) **que recibe una parte de la población** y el **tamaño de esa proporción de la población** (en porcentaje). En consecuencia, **esta forma de análisis** se centra en las **diferencias operativas en los datos en lugar de en un valor esperado**. Podemos relacionar este enfoque con la comprensión de los mercados Spot de EC2 y sus patrones de fijación de precios asociados. En particular, **podemos usar estos índices para detectar dos características**: **(i) ocurrencias de extremos de precio Spot**, y **(ii) la presencia de un componente estacional** (es decir, variaciones en los precios Spot que se deben a la hora del día) en un dado el mercado de instancia Spot.

**Sin embargo**, mientras que el **coeficiente de Gini y el índice de Theil pueden informarnos sobre la existencia de tales patrones** (únicamente a través de un único número cada uno), **no pueden revelar la ubicación y la forma de estos patrones**. Para comprender completamente la naturaleza del patrón de precio spot**, necesitamos un medio adicional: el gráfico de promedio móvil ponderado exponencial**. Este cuadro sirve como un método de detección de cambios en el que un **factor de olvido** A **controla la cantidad de información que se transmite desde el pasado hasta el presente**. **Para el análisis de los datos del precio Spot, un factor de olvido A = 0.5 tiene el mejor rendimiento**. En este nivel, los cambios en los precios Spot se reflejan con precisión en el gráfico del promedio móvil ponderado exponencial, incluida la magnitud de las fluctuaciones en cualquier dirección.

*Mercados invariables*: bastante sorprendente, n**o todos los mercados de instancias puntuales están sujetos a cambios** en su estructura de precios spot. Las razones de esta estabilidad son desconocidas. Solo podemos cuantificar el efecto. La naturaleza de los **índices de desigualdad asigna una puntuación de 0 a un conjunto de datos, si todas las instancias en este conjunto comparten la misma cantidad de un recurso dado**. Aplicado a los mercados de precios Spot esto significa que **un mercado en el que el precio Spot nunca cambia**, se le asigna un coeficiente de **Gini de 0**. El **mercado es invariable**. La relación RB (H) también es cero. Los mercados invariables no tienen un componente estacional. Podemos eliminar los mercados invariables del análisis, ya que no arrojan más información. Para Windows, hay 53 mercados invariables. Linux / UNIX y SUSE Linux forman parte de 21 mercados invariables cada uno. Sin embargo, es importante tener en cuenta que esta estabilidad solo se relaciona con el rendimiento pasado de los precios Spot. No excluye los cambios en el futuro, en particular, después del 19 de octubre de 2016, el final de nuestro período de observación.

*Mercados variables*: ¿Cuáles son los patrones típicos de precio Spot en los mercados de EC2? La Figura 1 ilustra la estructura del sistema operativo Windows. Los mercados de Linux / UNIX y SUSE Linux exhiben estructuras similares. Sin embargo, sin pérdida de generalidad, nos centraremos en las instancias de Windows solo en el resto de esta sección.

Del análisis de la Figura 1, descubrimos que las instancias de precio Spot se pueden categorizar en cuatro grupos con las siguientes características.

- **Gini bajo, RB bajo (II).** Los mercados al contado con estos tributos están ubicados cerca del origen del gráfico en la Figura 1. Destacamos las instancias de la zona A **Wa.m3.medium y Wa.a.micro**. Estos mercados son **independientes del tiempo** y **no disruptivos**. No existe un componente estacional en estos mercados y los precios al contado se distribuyen estrechamente en torno a la media (una característica señalada por un bajo coeficiente de Gini). Estos mercados prometen resultados altamente predecibles basados en los datos de precios históricos. Desde la perspectiva del consumidor EC2, estos mercados parecen relativamente seguros para pujar a una media histórica, aunque los resultados reales pueden fluctuar.

**- Gini bajo, RB alto (II).** Los mercados en este grupo poseen un fuerte componente estacional. Estos mercados **dependen del tiempo**, donde las fluctuaciones de los precios se limitan a horas específicas del día. Sin embargo, dado que el coeficiente de Gini es muy bajo, implica que las variaciones de precios son relativamente pequeñas y, por lo tanto, estos **mercados no son disruptivos**. Las instancias típicas en este grupo son **Wa.ml.nzediunt y Wb.mlanedium** con un valor para RB (II) de 49% y 33%, respectivamente. Las pujas avanzadas para estas instancias dependen en gran medida de la hora del día, aunque las diferencias pueden ser pequeñas. Es importante tener en cuenta que los valores de RB (11) entre 10% y 30% pueden ser un signo para cambiar los niveles de estacionalidad, como es el caso en el mercado Spot Wa.m3.xlarge. Esto significa que los mercados con estos valores de RB (II) pueden tener dos o más secciones con patrones de precios Spot estacionales, pero los precios Spot registrados varían sustancialmente entre esos componentes estacionales. Tal característica no necesariamente arroja coeficientes de Gini más altos, ya que los cambios de nivel estacional normalmente no implican picos extremos.

**- Alto Gini, bajo RB (II).** Los mercados al contado con altos coeficientes de Gini pueden ser muy perjudiciales. **Los valores de Gini superiores a 0.6 indican picos extremos ocasionales** en el historial de precios directos **que pueden ocurrir en cualquier momento durante el período de 24 horas**. Sin embargo, estos picos no son un fenómeno común en estos mercados. Estos mercados pueden arrojar resultados altamente predecibles, si los picos han sucedido mucho en el pasado. Un ejemplo de dicha instancia es **Wb.g2.8xlarge**. Mientras que **el precio Spot para este mercado se estableció en $ 0,8**, **hubo un período corto en julio de 2016, donde el precio Spot se registró en $ 38,7**. Los coeficientes de **Gini entre 0.2 y 0.6**, por otro lado, indican un grado de variabilidad **muy grande**. Los picos ocurren en una frecuencia más alta, pero pueden no ser tan extremos. Los mercados al contado con esta característica exhiben un componente local estacional, que puede obstaculizar el éxito de las ofertas.

**- Alto Gini, Alto RB (H).** Los mercados en este grupo exhiben un componente estacional emergente con picos frecuentes. Estos mercados son **independientes del tiempo y disruptivos**. Ejemplos de tales mercados son **Wa.c3.xlarge y Wb.c3.xlarge**. Una estrategia de licitación informada en estos mercados requiere una comprensión tanto de la estacionalidad como de la magnitud de las variaciones del precio Spot. Sin embargo, los valores que ambos índices pueden asumir al máximo están limitados por la naturaleza de los precios spot. Por el momento, los usuarios pueden pujar hasta 10 veces el precio On-Demand. Para que la proporción RB (11) del ratio "intermedio" supere el 50%, manteniendo un alto valor Gini, los precios spot extremos deben ubicarse en una hora específica del día. Tal característica no está presente en los datos históricos inspeccionados.

**Análisis de tendencia:**

La aplicación del **análisis econométrico de desigualdad** utilizando el **coeficiente Gini y la relación grupal** "intermedia" / 6 (11) **nos ha permitido agrupar los mercados EC2 Spot en diferentes grupos** (es decir, **identificar patrones típicos de precios spot**). Para caracterizar aún más estos patrones, es decir, **para identificar la forma asociada de estos patrones**, **inspeccionamos los gráficos de promedio móvil ponderado exponencial** de los mercados Spot.

El primer desafío surge de **encontrar un valor adecuado para el factor A de olvido**. La Figura 2 ilustra esto para el mercado Wb.ml.medium de Windows Spot. **En A = 1, EWMA se convierte simplemente en la media móvil durante todo el período de observación**. Existe un "efecto de suavizado" sustancial que distorsiona en gran medida la estructura real del patrón del precio spot. En general, **cuanto más cerca esté el valor de A a 1, mayor será la influencia de las observaciones pasadas**. Además, este alisamiento de curvas a menudo da como resultado una subestimación de los precios reales al contado. Por ejemplo, **entre el 2 de septiembre de 2016 y el 19 de octubre de 2016**, el **precio Spot** para **Wb.ml.medium** estuvo continuamente por **encima de $ 0.041**. Sin embargo, el EWMA en **A = 0.9** sugiere precios al contado **mucho más bajos**.

Podemos **reducir la influencia de los datos pasados ​​al establecer A más cerca de 0**. En A = 0, solo importa el valor actual (es decir, obtenemos un diagrama de dispersión en el que los precios Spot se representan con sus valores reales). A través de la experimentación, encontramos que **A = 0.5 produce un EWMA efectivo para los mercados Spot**. Los cambios en los precios Spot se reflejan con precisión en este valor del factor de olvido. Además, A = 0.5 evita tanto una subestimación como una sobreestimación de los precios Spot. Por lo tanto, argumentamos que una tabla EWMA **con un factor de olvido A = 0.5 proporciona una herramienta efectiva para identificar la forma de los patrones de precio spot,** lo que, además, también permite hacer inferencias sanas sobre las tendencias pasadas y futuras de la evolución del precio spot.

***Ejemplo de gráficas en distintos mercados***

La **Figura 3** representa los gráficos EWMA en **A = 0.5** para los mercados Spot con componentes estacionales. Los mercados Spot capturados a la izquierda de la Figura 3 (es decir, Wa.mlanedium y Wb.ml.medium) tienen bajos coeficientes de Gini. Los precios reales del Spot fluctúan dentro de un pequeño margen de $ 0.003. Sin embargo, la naturaleza periódica de estos cambios da lugar a un componente estacional definitivo. Una inspección detallada a través de un simple análisis de series de tiempo revela que los precios spot más altos se materializan regularmente entre las 12 a.m. y las 4 a.m. UTC en estos mercados.

Los mercados Spot que se muestran a la derecha de la Figura 3 (es decir, **Wb.c3.xlarge y Wa.c3.xlarge**), por otro lado, también están dotados de altos valores de Gini. Las variaciones en esos mercados ocurren con una magnitud mucho mayor (ap-prox. $ 5.00). Los cambios están asociados con períodos fijos, aunque ocurren en momentos diferentes. Encontramos que el mercado Spot Wa.c3.xlarge experimenta periodos de precios Spot altos a las 2 a.m. UTC, mientras que el mercado Spot Wa.c3.xlarge registra regularmente precios spot altos entre las 8 p.m. y 10 p.m. UTC. Una estrategia de oferta informada debe tener en cuenta la ubicación de los componentes estacionales. Los datos históricos sugieren que es posible asegurar un c2.xlarge de Windows en la región de Sydney por un precio spot promedio de $ 0.2074 hasta 20 horas, solo el 4% de los precios pico máximos. Cuando se ignora el componente estacional (o no se puede evitar de otro modo), los precios spot pueden saltar hasta $ 5,00.

Los perfiles para mercados Spot que carecen de un componente marino definitivo se muestran en la Figura 4. The Spot **mar-kets Wanz3.mediwn and Wa.t. 1 anicro**, a pesar de algunos picos registrados, también se caracterizan por bajos coeficientes de Gini. Estos mercados son más o menos estables, pero están sujetos a variaciones ocasionales en el precio spot que exceden el precio spot promedio en no más de una magnitud. Desde la perspectiva del usuario, estos mercados generalmente son seguros para ofertar en cualquier momento. Los gráficos EWMA brindan más pistas sobre el desempeño pasado y futuro de estos mercados. Por ejemplo, mientras que los precios al contado se mantuvieron estables y bajos, en una mediana de $ 0.0061, la historia más reciente revela una mayor actividad, que debe tenerse en cuenta mediante una estrategia de oferta informada.

El mercado **Wb.g2.8xlarge** proporciona un ejemplo con patrones extremos de precio Spot. Por más tiempo, el precio Spot registrado en este mercado es de $ 0,8. Sin embargo, en las dos primeras semanas de julio de 2016, el precio Spot para esta instancia fue de $ 38.7, 48 veces más alto que lo normal. El coeficiente de Gini para este mercado es 0.7476, lo que indica una dispersión muy grande en los precios Spot subyacentes. Por otro lado, el gráfico de EWMA revela que este pico pasó mucho tiempo atrás y no hay indicios de una recurrencia para el 19 de octubre de 2016, el final de nuestro período de análisis.

Finalmente, la Figura 5 ilustra los mercados Spot en los que se ha producido un componente estacional local (es decir, **Wa.m3.2xlarge**), o en los que los precios Spot se agregan en distintos niveles de componentes estacionales (es decir, Wa.m3.xlarge). El primero empujó el coeficiente de Gini más alto, mientras que el último redujo la relación de grupo "intermedio" RB (H). Ambos aspectos pueden complicar una estrategia de licitación informada. Sin embargo, un componente estacional definitivo da como resultado un mercado Spot altamente dependiente del tiempo. Por ejemplo, se debe tener especial cuidado para asegurar una instancia Wa.m3.xlarge. Los precios spot difieren significativamente en un período de 6 p.m. y 7 p.m. UTC.

1. ESTIMACIÓN DEL PRECIO DEL PUNTO

**En esta sección**, presentamos un modelo de **estimación del precio Spot** que explora el **historial de precios de Spot** para **ayudar** a los posibles **ofertantes** a **determinar** los **precios** de **oferta** **adecuados** en función de sus **necesidades específicas**. Para evaluar la efectividad de nuestro modelo propuesto, generamos 36,000 ofertas aleatorias por mercado Spot, totalizando 9.72 Ofertas Mio para los 270 mercados Spot analizados en la región de Sydney. Como resultado, obtuvimos 270 puntos de datos que capturaron 183 variables de estimación para la precisión de la estimación de nuestro método propuesto.

1. **Modelo de estimación**

La **ecuación** (1) muestra nuestra función de estimación del precio spot propuesta ***est (m, k, h, a),*** donde:

- **m - se refiere al mercado de instancia Spot que interesa al oferente,**

**- k - especifica la duración mínima de disponibilidad sostenida**, es decir, **la duración por la cual el postor desea acceso ininterrumpido a la instancia de spot,**

- **h - define la duración del período de retrospectiva en los datos históricos**, y

- **alpha- establece el nivel de confianza del precio de oferta estimado**.

El proceso inducido por est se puede entender de la siguiente manera:

"*Elija la oferta más pequeña del precio de oferta dentro del rango de los precios Spot registrados en el período de revisión de aspecto '****h****' para el mercado '****m****' de modo que la probabilidad de asegurar la instancia para '****k****' sea mayor o igual a '****alfa*** *'. "*

Por lo tanto, la **función de estimación est** **devuelve el menor precio de oferta** (dentro de los precios Spot reales SP (m, h)) que, de acuerdo con el historial h del mercado m, ha asegurado una instancia ininterrumpida durante k horas (spk (m, i) ) denota el precio Spot de referencia real) con una probabilidad alfa.

1. **Procedimiento de validación**

**¿Cuál es la utilidad de nuestra función de estimación?** Para probar su validez, diseñamos un escenario de **prueba** con las siguientes características:

- **Analizamos seis niveles de confianza esperados, a: 0.5, 0.6, 0.7, 0.8, 0.9 y 0.95**.

- **Definimos seis periodos de retrospectiva de historial de datos, h: 2, 7, 15, 30, 45 y 60 días**.

**- Seleccionamos cinco duraciones, k: 1, 2, 4, 8 y 16 horas.**

- Para todas las combinaciones axhxk, **seleccionamos aleatoriamente 200 puntos de inicio**, i, **entre el 2 de julio de 2016 y el 19 de octubre de 2016** (contabilizando también los márgenes requeridos de retrospectiva y duración) **y contamos el número de ofertas exitosas**, bidj ,,, ( m, i, k), que aseguran una instancia en el punto de partida i para una duración k en el mercado m con la probabilidad a y utilizando un período de retrospección h.

- **Realizamos la prueba para cada uno de los 270 mercados analizados EC2 Spot y registramos para cada mercado, m, una probabilidad de estimación west (m, k, h, a) = k) / 200**.

Para que nuestra función de estimación sea producir ofertas de precio Spot viables, requerimos que para cada mercado m, cresta (m, k, h, a)> a. Idealmente, aest (m, k, h, a) debería estar cerca de a para evitar la sobreestimación. Sin embargo, la utilidad real de est depende en gran medida de los patrones subyacentes del precio Spot en un mercado determinado, como se explica a continuación.

1. **La ejecución de la función estimación**

En las **tablas I y II** se muestra un **resumen** de alto nivel de los resultados de desempeño experimental para nuestro **método de estimación propuesto**. Encontramos que **para todos los niveles de confianza**, a, nuestro método de estimación arroja como promedio una estrategia de pujas que puede asegurar una instancia confiable al menos el 80% del tiempo. El rendimiento promedio de la estimación nunca cae por debajo del nivel a solicitado. Esto se aplica tanto al período de retrospectiva elegido como a la duración solicitada.

Un resumen de alto nivel puede informarnos sobre las características generales de nuestros métodos de estimación propuestos. Sin embargo, no puede capturar cualidades de rendimiento específicas del mercado. Para realizar un análisis específico del mercado, **debemos inspeccionar todos los pares posibles** (correlaciones) **de las variables de estimación**. Hay 183 variables de estimación que dan lugar a 16.653 pares individuales (es decir, coeficientes de correlación). Una herramienta efectiva para resumir e interpretar datos de este tamaño es el Análisis de componentes principales (PCA). **El objetivo de PCA es extraer conclusiones de las relaciones lineales entre variables al reducir los datos a las principales dimensiones de la variabilidad**. Más precisamente, PCA permite estudiar las semejanzas y diferencias entre las observaciones evaluadas desde un punto de vista multidimensional, y proporcionar un resumen descriptivo y la visualización del conjunto de datos subyacentes.

**PCA permite un análisis de rendimiento específico del mercado de nuestro método de estimación propuesto**. El **gráfico** que se muestra en la **Figura 6** captura los dos primeros componentes principales (es decir, el **primer plano principal**) de variabilidad en el rendimiento del estimador est con respecto a los mercados de precio Spot individuales. Los mercados spot se separan en función de su rendimiento de estimación respectivo. Los mercados con alto rendimiento de estimación (los valores de aest (m, k, h, a) están en el percentil 95) están mapeados a la derecha, mientras que los mercados con bajo rendimiento de estimación (los valores de aest (m, k, Ii, a) están en el percentil 25) están mapeados a la izquierda. Coincidentemente, la relación grupal "intermedia" RR (11) se correlaciona negativamente con el primer componente principal. Esto significa que los mercados Spot con componentes estacionales fuertes se asignan a la izquierda, y los mercados Spot que carecen de un componente estacional se asignan a la derecha. Es decir, el rendimiento del estimador es una función del componente estacional.

El **segundo componente principal (eje y en la Figura 6**) separa el rendimiento del mercado por su coeficiente de Gini respectivo y las combinaciones a x h x k. Los mercados en los cuadrantes superiores tienen coeficientes de Gini más altos que los de los cuadrantes inferiores. Los mercados mapeados en el cuadrante inferior exhiben un mejor rendimiento de estimación en valores más pequeños de hyk, mientras que los mercados mapeados en los cuadrantes superiores se desempeñan mejor en valores más altos de h y k.

Lo más importante, **PCA permite un análisis de similaridad**. Los puntos de datos que están mapeados uno cerca del otro son similares. **PCA separa claramente los mercados de Windows y Linux / UNIX de los mercados de SUSE Linux**. SUSE Linux mar-kets tiene un componente estacional muy débil, si es que tiene alguno, y los precios spot se manifiestan dentro de márgenes muy pequeños. Es por esta razón que la mayoría de los mercados de SUSE Linux están mapeados en el cuadrante inferior. También significa que la estimación del precio de oferta en los mercados de SUSE Linux es más confiable en los valores pequeños de h y k. Los mercados de Windows y Linux / UNIX, por otro lado, exhiben patrones comparables, incluyendo picos extremos ocasionales y fuertes componentes estacionales.

Sin embargo, su mapeo predominante al cuadrante superior significa que la estimación del precio de oferta es más confiable en valores más grandes de hyk.

Finalmente, la Figura 7 proporciona un resumen del rendimiento de la estimación de los mercados de Windows seleccionados discutidos en la sección anterior. Las asignaciones concuerdan con los coeficientes de Gini observados y las proporciones "intermedias" RB (H). Los mercados con un fuerte componente estacional (altos valores de RB (H)) son más difíciles de estimar. Los valores altos de Gini, por otro lado, pueden no impedir la estimación, como es el caso del mercado Wb.g2.8xlarge. Este mercado exhibe un excelente rendimiento de estimación. Por lo tanto, está asignado a la extrema derecha en el primer plano principal. Por otro lado, los picos perturbadores en los precios al contado en agosto de 2016 para el mercado Wa.rn3.2xlarge (ver Figura 5a) hacen que la estimación sea más confiable de los valores altos de hyk.

1. TRABAJO RELACIONADO

El mercado spot de EC2 fue anunciado por primera vez por Amazon el 14 de diciembre de 2009 [13]. Si bien el precio de la instancia puntual se estableció inicialmente por región, se introdujo un precio de instancia puntual por zona de disponibilidad en 2011 [14]. En 2013, Amazon introdujo un límite de precio de oferta predeterminado que se estableció en 4 veces el precio On Demand [15]. Desde entonces se ha restablecido a 10 veces el precio a pedido [16]. Más recientemente, Amazon ha presentado la Spot Fleet y el Spot Bid Advisor [5]. Spot Bid Advisor es una herramienta basada en web que presenta a los consumidores de EC2 estadísticas relevantes en múltiples zonas de disponibilidad sobre la volatilidad de los precios y les ayuda a tomar mejores decisiones de compra.

Uno de los primeros análisis del mercado spot de Amazon EC2 se presentó en [3], donde los autores estudiaron la dinámica del mercado y llegaron a la conclusión de que los precios Spot no están totalmente orientados al mercado sino a un precio de reserva oculto. Sin embargo, los mismos autores actualizaron posteriormente sus hallazgos en [17] y declararon que las características del mercado spot EC2 habían cambiado para reflejar una asignación impulsada por el mercado. Del mismo modo, en [18], los autores han proporcionado un análisis estadístico de las instancias puntuales en términos del precio spot y el tiempo entre los precios, es decir, el tiempo entre los cambios de precios. En su enfoque, la dinámica del precio Spot y el tiempo entre precios se modela como una mezcla de distribución gaussiana que se puede usar en el diseño de algoritmos de programación estocástica y mecanismos tolerantes a fallas.

Hay una serie de estudios que han analizado los precios del mercado Spot para desarrollar estrategias de oferta que garanticen un cierto nivel de garantías de rendimiento al tiempo que minimizan los costos de uso [19] [20]. Por ejemplo, en [19], los autores han investigado las estrategias de pujas óptimas para trabajos paralelos de MapReduce en dos configuraciones: solicitudes únicas y solicitudes persistentes. En [20], los autores han estudiado estrategias de oferta óptimas para trabajos con fechas límite duras. Del mismo modo, en [21], los autores diseñaron una estrategia de oferta desde la perspectiva de un bróker. Sin embargo, solo han considerado un mercado único en lugar de buscar alternativas en otros mercados al contado.

En los autores han argumentado que no hay necesidad de estrategias de licitación sofisticadas. En primer lugar, la naturaleza del mercado Spot es tal que existe una amplia gama de precios de oferta que producirán el costo óptico y la disponibilidad. En segundo lugar, como hay un gran número de mercados al contado simultáneos, siempre hay un mercado con recursos excedentes disponibles. Por lo tanto, en lugar de centrarse en estrategias de pujas complejas, los usuarios pueden adoptar estrategias de pujas triviales y centrarse en modificar las aplicaciones para que puedan migrarse fácilmente de una región a otra. Este es uno de los primeros estudios que considera varios mercados Spot, aunque no se centra en la comprensión de la dinámica del mercado Spot, sino que se centra en cómo cambiar de un mercado a otro cuando cambia el precio Spot.

En, los autores han presentado una metodología para predecir el valor de la oferta para garantizar que la duración de la solicitud excede una duración fija con una probabilidad dada. Utilizaron los requisitos de la aplicación HPC como motivación para su investigación y utilizaron un método de análisis de series temporales llamado Queue Bounds Estimation from Time Series (QBETS) para predecir el límite superior en el precio de oferta y el límite inferior en la duración.

En [23], los autores se han centrado en el tiempo útil del servidor de una Instancia puntual y argumentaron que las variaciones en el precio Spot afectan el valor inherente del tiempo del servidor comprado en el mercado. Han presentado el argumento de que a medida que más usuarios aprovechan las oportunidades de arbitraje para reducir los costos de uso, más disminuye el valor de las instancias de Spot. Como alternativa, han propuesto el concepto de garantías transitorias donde proveedores como Amazon EC2 y Google Compute Engine ofrecen capacidad inactiva por un precio fijo pero con una garantía transitoria en su tiempo medio de revocación. Ofrecer garantías transitorias proporciona a los usuarios garantías estadísticas sobre su disponibilidad, volatilidad y previsibilidad. [24] propone un nuevo modelo para analizar el historial de precios Spot basado en cuatro características clave: duración de la instancia puntual, precio spot promedio durante la vida útil, revocaciones simultáneas y retraso de inicio. Una diferenciación clave de este trabajo y otros trabajos es el enfoque en las revocaciones simultáneas y el retraso de inicio.

En [25], los autores han propuesto un servicio llamado Spot-light que utiliza sondeos activos para estudiar la dinámica interna de los mercados de Amazon EC2 (On-Demand y Spot). Han basado su enfoque en la idea de que los precios en tiempo real en los mercados al contado se correlacionan de forma vaga con el suministro y la disponibilidad de las instancias On-Demand. La principal motivación para el enfoque propuesto es obtener algunos conocimientos sobre la disponibilidad de On-Demand y las instancias puntuales, ya que Amazon no proporciona información. El documento ha propuesto dos niveles de sondeo: (a) indagar sobre la disponibilidad de instancias On-Demand siempre que el precio Spot correspondiente suba, y (b) buscar instancias Spot en diferentes niveles de oferta. El sondeo es consciente de los costos (ya que la presentación de solicitudes a pedido tiene un costo relacionado) y admite el sondeo con la familia de instancias y entre las zonas de disponibilidad. Obtener mejores conocimientos sobre la disponibilidad de instancias en los diferentes mercados puede ayudar a hacer una selección más óptima de las instancias y opciones de compra.

Diferenciación: el objetivo principal de nuestro trabajo propuesto es comprender la dinámica de fijación de precios en los mercados Spot de Amazon EC2 basándose en los datos históricos y estimar las ofertas mínimas en función de los requisitos. En lugar de utilizar los enfoques convencionales de las estadísticas de tendencia central para estudiar la volatilidad, utilizamos índices de desigualdad para explicar el comportamiento del mercado Spot de EC2. Hasta donde tenemos conocimiento, nuestro enfoque basado en índices de desigualdad es el primero en cuantificar la volatilidad de los mercados Spot y calificar el riesgo asociado con las ofertas en ellos. Un beneficio clave de nuestro enfoque es que puede resumir todo el historial de precios de Spot utilizando tres métricas y usarlas para categorizar los mercados spot en diferentes clusters basados en los dos índices de desigualdad. Otra diferencia clave es que no limitamos nuestro análisis a mercados únicos y permitimos la comparación directa de la dinámica de fijación de precios con y entre los mercados.

1. CONCLUSIÓN

En este documento, presentamos un enfoque estadístico descriptivo para estimar precios mínimos de oferta que cumplen con los requisitos de costo y disponibilidad basados ​​en información histórica de precios. Utilizamos tres medidas estadísticas, el coeficiente de Gini, el índice de Theil y la media móvil ponderada exponencial, para detectar y localizar patrones de fijación de precios típicos, incluidos picos extremos, componentes estacionales y tendencias en los precios al contado. También propusimos una fórmula de estimación de ofertas mínima que explora la información histórica de precios Spot para ayudar a los licitantes a tomar decisiones informadas sobre las ofertas. Los resultados experimentales demuestran que la función de estimación del precio de oferta propuesta es una función implícita tanto de la presencia de componentes estacionales (capturados por el índice de Theil) como de la distribución de picos extremos (capturados por el coeficiente de Gini). Además, nuestra estimación arroja, en promedio, una estrategia de oferta que puede asegurar de forma confiable una instancia al menos el 80% del tiempo con una garantía mínima entre 50% y 95%.

Como trabajo futuro, tenemos la intención de ampliar la función de estimación de ofertas para hacer estimaciones sensibles al tiempo, es decir, además de estimar el precio mínimo de oferta, también se estima el mejor tiempo de oferta. Además, nos gustaría hacer que la función de estimación de ofertas sea más adaptable, es decir, el período de retrospección se puede seleccionar automáticamente en función de los valores correspondientes de Gini y RB (II).