

Apoyo a la Toma de Decisiones en la Compra de IaaS*

Octavio Martín-Díaz, Pablo Fernández, and Antonio Ruiz-Cortés

Dpto. Lenguajes y Sistemas Informáticos
ETS. Ingeniería Informática – Universidad de Sevilla
41012 Sevilla, España – Spain
omartindiaz@us.es, pablofm@us.es, aruiz@us.es

Resumen. La dificultad para decidir la compra de un IaaS (*Infrastructure as a Service*) depende de la complejidad de las opciones de compra dadas por su proveedor y de la complejidad del plan del cliente que quiere realizarla. Es habitual que estos tipos de servicios ofrezcan muchas configuraciones de uso diferentes, y para cada una de ellas sea posible disponer de varias opciones de compra. De este modo, decidir la mejor compra se convierte en una tarea que consume mucho tiempo, tediosa y propensa a errores. En este trabajo inicial, caracterizamos el problema con un caso de estudio ilustrativo y presentamos los desafíos inmediatos para mejorar las herramientas de soporte actualmente disponibles.

Palabras claves: IaaS, Opciones de Compra, SLA

1. Introducción y Motivación

La compra de un IaaS permite adquirir recursos de procesamiento y almacenamiento como alternativa a la adquisición de una infraestructura informática propia. Esta alternativa permite abaratar los costes asociados a la creación de dichas infraestructuras y su posterior renovación. El escenario de compra típico puede ser el siguiente: (1) un proveedor oferta las *opciones de compra* para un servicio, las cuáles suelen estar fijadas para una configuración determinada del servicio y un tiempo de uso; y (2) un cliente presenta su *plan de uso* y selecciona a partir de las opciones de compra un *plan de compras* que le sea rentable.

La dificultad para obtener un plan de compra depende de la complejidad de las opciones de compra y también de la complejidad del plan de uso del cliente que quiere hacer dicha compra. Las opciones de compra de un IaaS por parte de un proveedor pueden estar ligadas a servicios altamente configurables, de manera que la lista de precios puede llegar a ser abrumadora.

* Este trabajo ha sido cofinanciado por el Gobierno de España por el proyecto TAPAS (TIN2012-32273) y la Red Temática Científico-Tecnológica en Ciencias de los Servicios (TIN2011-15497-E), y la Junta de Andalucía por los proyectos THEOS (TIC-5906) y COPAS (P12-TIC-1867).

Por ejemplo, Amazon EC2 ofrece actualmente 7954 configuraciones de instancia diferentes [5]. Además, tanto configuraciones como opciones de compra pueden estar sujetas a variabilidad, es decir, el proveedor puede añadir o quitar configuraciones, añadir o quitar parámetros de las propias configuraciones, añadir o quitar opciones de compra, cambiar precios, o incluso cambiar el propio modelo subyacente para las opciones de compra, por ejemplo, modificando el método para calcular los precios.

Respecto al cliente, los planes de uso pueden referirse a uno o más servicios, del mismo o diferente tipo, con la misma o diferente configuración. Por ejemplo, un entorno de virtualización de prácticas de Ingeniería del Software tiene que atender a diferentes asignaturas, con diferente software, diferentes grupos y horarios, que pueden estar solapados.

Teniendo en cuenta las opciones de compra y un plan de usuario, el objetivo es encontrar un plan de compra óptimo, consistente en conocer cuántas instancias de cada opción de compra hay que reservar para ejecutar el software requerido sobre el IaaS.

En la actualidad existen herramientas que permiten la búsqueda óptima de un tipo de instancia entre varios proveedores [4], o varios tipos de instancia independientes [3]. Sin embargo, la toma de decisión sobre opciones de compra en escenarios complejos, tal y como describimos, sigue siendo un proceso manual que consume tiempo, es tedioso y propenso a errores. Así pues, se hace necesario un conjunto de técnicas y herramientas que soporte de manera automatizada el proceso de compra.

En este trabajo inicial, caracterizamos el problema y argumentamos los desafíos inmediatos, poniendo de relieve los parámetros clave mediante un caso de estudio con opciones de compra altamente configurables y planes de uso complejos.

El resto del trabajo se organiza como sigue. La sección 2 detalla un caso de estudio ilustrativo del escenario para caracterizar el problema. A continuación, la sección 3 presenta los desafíos para abordar el problema y la sección 4 el trabajo relacionado. Finalmente, la sección 5 expone nuestras principales conclusiones y trabajo futuro.

2. Caso de Estudio

En esta sección presentamos un ejemplo ilustrativo para identificar los parámetros clave en el escenario de compra de IaaS: servicios altamente configurables y complejidad de los planes de uso.

2.1. Opciones de Compra en Amazon EC2

Actualmente Amazon EC2 es uno de los IaaS más populares. Uno de sus elementos clave es la *instancia*, que representa un recurso virtualizado de procesamiento y memoria sobre el que se despliega una máquina virtual. La configuración de una instancia incluye un conjunto de parámetros muy variado.

Entre otros parámetros:

- Nombre identificativo de la instancia, por ejemplo, “m3.medium” o “m3.large”.
- Uso previsto de la instancia: *light*, *medium* o *heavy*.
- Parámetros de procesamiento/memoria: número de CPU, 32 ó 64 bits, RAM o almacenamiento SSD/HD.
- Volumen de entrada y salida de datos: *very low*, *low*, *moderate*, *high* o *10-gigabit*.
- Duración de 1 ó 3 años.
- Sistema operativo: Linux, Windows, SQL-Windows o SQL-Web-Windows.

Cada configuración de instancia tiene sus opciones de compra (*purchasing options*) cuyos precios podrán variar según la zona en la que se encuentre desplegada la infraestructura real que le da soporte. Destacamos principalmente:

- *Bajo demanda* Se paga cuando se usa. Es la opción más cara, solamente recomendada para un uso esporádico de la instancia.
- *Reserva* Se paga por adelantado una parte del precio (*up-front*) a cambio de disminuir el pago por hora y garantizar la disponibilidad de la instancia. Hay tres opciones, a 1 y 3 años:
 - Ligero (*light*) Opción recomendada cuando se prevee que la instancia tenga un uso no intensivo, i.e. pocas horas cada día. El precio es muy similar (o igual en algunos casos) al uso bajo demanda, pero la disponibilidad de la instancia esta garantizada. Se cobran las horas de uso efectivo.

The screenshot shows the Amazon EC2 pricing calculator interface. At the top, it displays the estimated monthly bill of \$25.62. The configuration includes one m3.medium instance in the Europe (Ireland) region, running Linux on m3.medium. The billing option is set to 'On-Demand (No Contract)'. The table below shows the pricing for various options:

Select	Name	Hourly Price	Upfront Price	1 Year Cost	3 Year Cost	Effective Monthly Cost
<input type="radio"/>	On-Demand (No Contract)	0.077	---	676.44	2029.32	56.37
<input type="radio"/>	1 Yr Light Reserved	0.077	110.00	786.44	2359.32	65.54
<input checked="" type="radio"/>	1 Yr Medium Reserved	0.035	101.00	408.44	1465.32	40.71
<input type="radio"/>	1 Yr Heavy Reserved	0.028	222.00	468.00	1404.00	39.00
<input type="radio"/>	3 Yr Light Reserved	0.043	172.00	---	1832.32	50.90
<input type="radio"/>	3 Yr Medium Reserved	0.029	286.00	---	1050.28	29.18
<input checked="" type="radio"/>	3 Yr Heavy Reserved	0.023	337.00	---	943.24	26.21

Figura 1. Calculadora de Amazon

- Medio (*medium*) Opción recomendada cuando se prevee un uso más continuado respecto a la anterior, de manera que el pago por adelantado es mayor y el precio/hora menor. Se cobran las horas de uso efectivo.
- Intenso (*heavy*) Opción recomendada cuando se prevee un uso casi continuo de la instancia. Adelanto de mayor cuantía, pero al precio/hora más bajo. En este caso se cobran todas las horas que cubren el período reservado, se haya hecho uso efectivo o no.

Dado que el precio de cada opción de compra puede variar en función de la configuración elegida, el número de instancias y la intensidad prevista de uso, encontrar el coste del servicio se convierte en una actividad tediosa y propensa a errores. Más si cabe cuando se agregan diferentes tipos de instancia u otras opciones que (por simplicidad) no hemos tenido en cuenta en este trabajo, entre ellas parámetros sobre volúmenes y transferencia de datos, balanceo de carga o IP elásticas.

Amazon ofrece una calculadora web para obtener dicho coste e incluso indicar la opción de compra recomendada. Por ejemplo, la Fig. 1 muestra su aspecto para una instancia de Amazon EC2 “*m3.medium*” de propósito general, con 1 CPU virtual, equivalente a 3 unidades de computación EC2, 3.75Gb RAM, 4Gb SSD, con Linux, en la Unión Europea¹. Nótese que, para reservas de 1 año de duración, el porcentaje de uso debe ser superior al 19% (unas 1650 horas anuales) para que la reserva de uso medio sea más rentable que la opción ligera, y superior al 80% (unas 7000 horas anuales) para que la reserva de uso intenso sea la más rentable.

2.2. Planes de Uso de un Entorno de Virtualización de Prácticas

Presentamos un caso de estudio sobre *virtualización del ecosistema de herramientas de prácticas* en la docencia de Ingeniería del Software. Estas prácticas pueden necesitar una evolución muy dinámica desde dos puntos de vista. Desde el primero, las prácticas suelen utilizar software que requiere de recursos crecientes, siendo más que probable que queden obsoletos en un relativo corto período de tiempo. Por el segundo, el número variable de estudiantes realizando prácticas a lo largo de un curso lectivo. Por ello, la virtualización hecha posible por la compra de una IaaS ofrece perspectivas tanto de flexibilidad como de ahorro frente a dicha necesidad.

En Amazon EC2, la virtualización se consigue mediante la reserva de instancias, cuyas configuraciones sean adecuadas, sobre las que se despliega el ecosistema. Para ello, el profesor de una asignatura debe indicar:

- La configuración mínima necesaria de la instancia para ejecutar el ecosistema.
- La planificación de prácticas a lo largo del curso, que incluye fundamentalmente:

¹ <http://calculator.s3.amazonaws.com/index.html>

- El horario semanal de clase de los grupos de prácticas, así como el número de alumnos matriculados en cada grupo.
- Las fechas y horas de exámenes prácticos.
- Las aulas de libre acceso, proporcionando el número estimado de estudiantes que quieran utilizarlas y los períodos de mayor demanda.

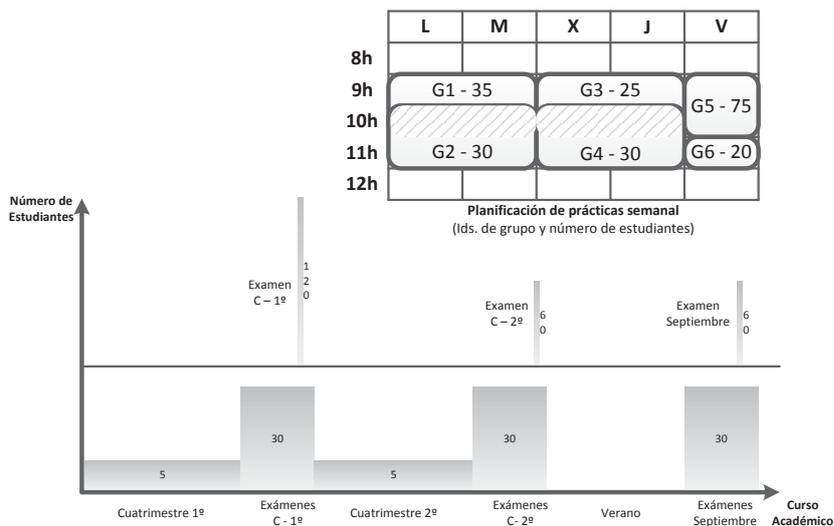


Figura 2. Ejemplo de planificación de prácticas

A modo de ejemplo, véase la planificación de la Fig. 2. Nótese que el calendario lectivo del curso académico consta de dos períodos cuatrimestrales, donde se imparten clases prácticas con una periodicidad semanal:

- Hay 4 grupos que reparten sus 4 horas semanales en dos días, existiendo solapes en sus horarios; más otros 2 grupos con menor número de horas. Estos grupos tienen matriculados a 35, 30, 25, 30, 20 y 75 estudiantes, respectivamente, para totalizar 215 estudiantes.
- Hay 3 períodos de exámenes, con un día previsto para las pruebas prácticas de la asignatura. Se prevee que haya una asistencia de 120 estudiantes en la convocatoria del primer cuatrimestre y 60 en la convocatoria del segundo cuatrimestre o en septiembre. Todas ellas tienen una duración de 4 horas.
- El aula abierta tiene prevista una asistencia de 5 estudiantes cualquier día durante el curso lectivo por 10 horas diarias. En períodos de exámenes y dos semanas previas, la asistencia puede ascender hasta 30 estudiantes.
- No hay docencia durante el período de vacaciones de verano².

² Por simplicidad, se han obviado otros tipos de descanso vacacionales o días festivos.

2.3. Obtención del Plan de Compras

Una vez que conocemos las opciones de compra que oferta Amazon EC2 y el plan de uso del profesor de una asignatura, es el momento de determinar el plan de compras, que consiste en obtener el número y tipo de instancias a reservar, que sea la opción más económica.

Services Estimate of your Monthly Bill (\$ 327.10)

Choose region: Europe (Ireland) Inbound Data Transfer is Free and Outbound Data Transfer is 1 GB free per region per month

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers. Amazon Elastic Block Store (EBS) provides persistent storage to Amazon EC2 instances. [Clear Form](#)

Compute: Amazon EC2 Instances:

Description	Instances	Usage	Type	Billing Option	Monthly Cost
G1	35	4 Hours/Week	Linux on m3.medium	1 Yr Light Reserved	\$ 48.65
G2	30	4 Hours/Week	Linux on m3.medium	1 Yr Light Reserved	\$ 41.70
G3	25	4 Hours/Week	Linux on m3.medium	1 Yr Light Reserved	\$ 34.75
G4	30	4 Hours/Week	Linux on m3.medium	1 Yr Light Reserved	\$ 41.70
G5	75	2 Hours/Week	Linux on m3.medium	1 Yr Light Reserved	\$ 52.50
G6	20	1 Hours/Week	Linux on m3.medium	1 Yr Light Reserved	\$ 7.80

Add New Row

(a) Necesidades para los grupos de prácticas por separado

Services Estimate of your Monthly Bill (\$ 318.35)

Choose region: Europe (Ireland) Inbound Data Transfer is Free and Outbound Data Transfer is 1 GB free per region per month

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers. Amazon Elastic Block Store (EBS) provides persistent storage to Amazon EC2 instances. [Clear Form](#)

Compute: Amazon EC2 Instances:

Description	Instances	Usage	Type	Billing Option	Monthly Cost
G20	20	15 Hours/Week	Linux on m3.medium	1 Yr Light Reserved	\$ 100.20
G25	5	14 Hours/Week	Linux on m3.medium	1 Yr Light Reserved	\$ 23.50
G30	5	12 Hours/Week	Linux on m3.medium	1 Yr Light Reserved	\$ 20.05
G35	5	8 Hours/Week	Linux on m3.medium	1 Yr Light Reserved	\$ 13.50
G55	20	6 Hours/Week	Linux on m3.medium	1 Yr Light Reserved	\$ 40.20
G65	10	4 Hours/Week	Linux on m3.medium	1 Yr Light Reserved	\$ 13.90
G75	10	2 Hours/Week	Linux on m3.medium	1 Yr Light Reserved	\$ 7.00

Add New Row

(b) Necesidades ajustadas para los grupos de prácticas

Figura 3. Comparativa de Estrategias de Compra

A modo de ejemplo y por simplicidad, tengamos en cuenta solamente los 6 grupos de estudiantes. El plan de compra más simple puede consistir en reservar una instancia de uso ligero para cada estudiante, es decir, para el G1 con 35 estudiantes, se reservan 35 instancias, y así sucesivamente. Entonces, la calculadora de Amazon nos da un coste de 327,10\$ mensuales, como aparece en la Fig. 3(a), con la configuración ya vista anteriormente.

Si estudiamos en detalle el plan de uso, podemos encontrar planes de compra más eficientes y económicos. Por ejemplo, teniendo en cuenta los solapes en los horarios y que no siempre todos los estudiantes están haciendo prácticas a la vez. La Fig. 4 ilustra el método para conocer cuántos estudiantes simultáneos hay, durante cuántas horas, a lo largo de la semana.

La Fig 3(b) muestra el plan de compra resultante, cuyo coste baja a 318,35\$ mensuales. Este plan de compra más ajustado se interpreta como sigue: entre los diferentes grupos de estudiantes, hay 20 estudiantes haciendo prácticas simultáneamente durante 15 horas, otros 5 más haciendo prácticas durante 14 horas, etc. hasta que llegamos a un ultimo grupo de 10 estudiantes más durante 2 horas, que cubre el pico de estudiantes que queda en el G5.



Figura 4. Cómputo de horas totales en el plan de uso

En este escenario, cabe preguntarse si acumulando horas de los planes de uso de diferentes asignaturas, aunque con necesidades de configuración similares, la reserva de instancias de uso medio o intensa llega a ser más económica. Estos cálculos son difíciles y necesitan ser automatizados, pero hasta lo que sabemos, Amazon no ofrece API de su calculadora y tampoco soporta el coste de varios grupos de instancias multiplexables en el tiempo.

3. Desafíos

La toma de decisión para seleccionar la mejor opción de compra en escenarios con servicios altamente configurables y planes de uso complejos ofrece un conjunto de desafíos que plantean un camino de investigación prometedor:

- Automatizar el proceso de compra a partir de documentos de acuerdos de nivel de servicio. Por ejemplo, se puede utilizar WS-Agreement [1], que es el estándar *de facto* para especificar acuerdos de nivel de servicio en la actualidad. Estos acuerdos podrían ser utilizados como documentos que contienen la información necesaria para expresar las opciones de compra (mediante plantillas) o los planes de uso (mediante ofertas de acuerdo).

- Mayor complejidad en los planes de uso. El soporte para expresar el plan de uso que ofrece Amazon u otros proveedores es limitado para algunos escenarios, lo que lleva a:
 - Que sean necesarios algunos cálculos extra para calcular el porcentaje de uso, debido por ejemplo a que los períodos se expresen en el mejor de los casos indicando el número de horas por día: de lunes a viernes de 9 a 15 horas, excepto los viernes, hasta las 14 horas. Los tiempos de uso debieran estar por tanto en términos del dominio del problema.
 - Que en ocasiones, lo que es más grave, se opte por un plan de compras que no es el óptimo, por ejemplo cuando no se tiene en cuenta la posibilidad de que los grupos compartan instancias del servicio cuando ello es posible. Los grupos de instancias deben ser multiplexables en el tiempo, es decir, hay que tener en cuenta los solapes y que el tamaño de los grupos para un mismo tipo de instancia sea variable en el tiempo.
 - Que los tipos de instancia puedan ser homogéneos o heterogéneos, es decir, que se necesitan instancias con diferentes configuraciones.

Las cuestiones relativas a la temporalidad ya han sido estudiadas en el contexto de WS-Agreement en [10].

- La Fig. 5 muestra la familia de *funciones de puntos de equilibrio* asociadas a cada opción de compra para 1 año correspondiente a la configuración “*m3.medium*”, presentada en la sección 2. Nótese que se ha destacado la envolvente en la familia de curvas que determina el coste mínimo. El análisis de las envolventes muestra los umbrales sobre porcentajes de uso que determinan las opciones de compra más económicas, de manera que se puede interpretar el cálculo de coste más económico como el valor de la envolvente para el porcentaje de uso indicado.

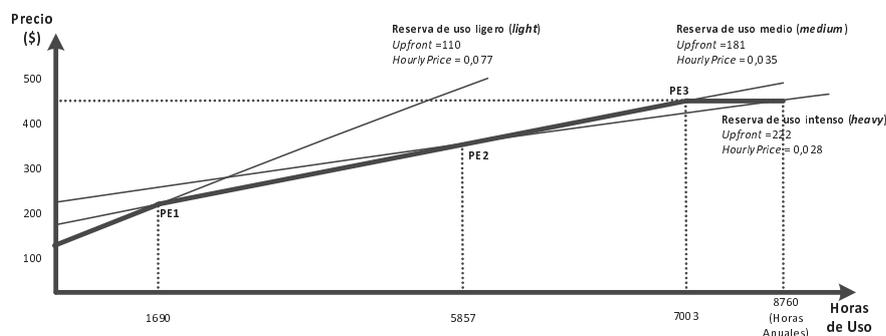


Figura 5. Familia de Puntos de Equilibrio

Esta interpretación facilitará en gran medida la automatización del cálculo del coste y se protegerá de las variaciones de las diferentes políticas de compra siempre que sea posible obtener automáticamente dicha envolvente a partir de la información disponible.

- Protección frente a las variaciones de las opciones de compra ofertadas por los proveedores. Además, diferentes proveedores pueden ofertar con diferentes modelos de compra; así mismo, un mismo proveedor puede modificar tanto el modelo como las opciones a lo largo del tiempo.
- Considerar otros modelos de compra. Las ofertas de los proveedores de IaaS actuales incluyen otros tipos ofertas dinámicas; por ejemplo, en Amazon EC2 además de las ofertas comentadas en nuestro caso de estudio, existen otros dos mercados adicionales: el mercado secundarios de instancias reservadas y el mercado de instancias de uso puntual (*Spot instances*).

4. Trabajo Relacionado

Hasta lo que conocemos, no hemos encontrado una propuesta que afronte el problema de la compra de servicios de IaaS altamente configurables con planes de uso complejos por parte del cliente. Pero sí que hay propuestas para solucionar partes más simples del mismo:

- En [5] se propone un método para encontrar la configuración óptima para una instancia, entre las miles de configuraciones que oferta Amazon EC2, en el contexto de migración de las aplicaciones software a la IaaS. En nuestro caso de estudio, este método puede utilizarse para encontrar la configuración mínima necesaria para el entorno de virtualización del ecosistema.
- En [4] se propone una herramienta que permite encontrar el mejor precio entre diferentes proveedores para una configuración dada. En nuestro caso de estudio, esta herramienta puede utilizarse para encontrar el proveedor óptimo para una configuración.
- En [3] se propone una herramienta similar, que permite encontrar el mejor precio entre diferentes proveedores para múltiples configuraciones de instancia independientes. Como la herramienta anterior, puede utilizarse en nuestro caso de estudio para encontrar el proveedor óptimo.

Entre las propuestas IaaS que tienen en cuenta el aspecto de la temporalidad en contexto similares al nuestro, encontramos:

- [2] propone el establecimiento automático de acuerdos de nivel de servicios, basado en WS-POLICY, con un caso de estudio basado en Amazon EC2.
- [7] propone el despliegue óptimo en una IaaS de un conjunto de servicios, que tiene en cuenta la variación del número de usuarios y el patrón probabilístico de invocaciones a dichos servicios.
- [11] propone algunos esquemas de negociación en IaaS. Presenta un caso de estudio en Amazon EC2 y propone un modelo de preferencias basado en funciones de descuentos y sensibles a la temporalidad.
- [8] también propone el problema de la negociación de SLA en IaaS.
- [9] analiza el problema del pago excesivo en IaaS debido al uso intensivo de servicios durante un periodo temporal, bajo un esquema pre-pago de servicios.
- [6] afronta el problema de analizar el uso concreto de las licencias software en IaaS, dado el gran dinamismo que hay en este contexto.

5. Conclusiones y Trabajo Futuro

En este trabajo inicial hemos presentado una caracterización al problema la toma de decisión para seleccionar la mejor opción de compra en escenarios con servicios altamente configurables y planes de uso complejos. Como caso de estudio se ha analizado el servicio de Elastic Computing Cloud (EC2) proporcionado por Amazon para la virtualización de entornos de prácticas. A partir de este análisis se han identificado los parámetros clave del problema que puedan servir como base para el desarrollo de sistemas de apoyo a la toma de decisiones en este contexto, exponiéndolos como una serie de desafíos abiertos al futuro inmediato.

Referencias

- [1] A. Andrieux, K. Czakowski, A. Dan, K. Keahey, H. Ludwig, T. Nakata, J. Pruyne, J. Rofrano, S. Tuecke, and M. Xu. Web Services Agreement Specification (WS-Agreement) Version 1.1 draft 20, September 2006.
- [2] M. Chhetri, Q. Bao Vo, and R. Kowalczyk. Policy-based Automation of SLA Establishment for Cloud Computing Services. In *12nd Intl. Symp. on Cluster, Cloud, and Grid Computing*, pages 164–171, Ottawa, Canada, 2012.
- [3] Clouddorado.com. Cloud Computing Price Comparison Engine. <http://www.clouddorado.com/>.
- [4] CloudScreener.com. Cloud Computing Comparison and Evaluation. <http://www.cloudscreener.com/>.
- [5] J. García Galán, O. Rana, P. Trinidad, and A. Ruiz-Cortés. Migrating to the Cloud: a Software Product Line based Analysis. In *3rd Intl. Conf. on Cloud Computing and Service Science*, pages 416–426, Aachen, Germany, 2013.
- [6] M. Kim, H. Chen, J. Munson, and H. Lei. Management-Based Licensing Discovery for the Cloud. In *10th Intl. Conf. on Service-Oriented Computing*, pages 499–506, Shanghai, China, 2012.
- [7] K. Konstanteli, T. Cucinotta, K. Psychas, and T. Varvarigou. Admission Control for Elastic Cloud Services. In *5th Intl. Conf. on Cloud Computing*, pages 41–48, Honolulu, Hawaii, 2012.
- [8] K. Lu, R. Yahyapour, E. Yaqub, and C. Kotsokalis. Structural Optimization of Reduced Ordered Binary Decision Diagrams for SLA Negotiation in IaaS of Cloud Computing. In *10th Intl. Conf. on Service-Oriented Computing*, pages 268–282, Shanghai, China, 2012.
- [9] L. Lu and E. Elmroth. A Time-Interval-Based Credit Rerservation Approach for Prepaid Composite Services in Cloud Environments. In *9th European Conf. on Web Services*, pages 158–165, Lugano, Switzerland, 2011.
- [10] C. Müller, O. Martín-Díaz, A. Ruiz-Cortés, M. Resinas, and P. Fernández. Improving Temporal-Awareness of WS-Agreement. In *5th Intl. Conf. on Service-Oriented Computing*, volume 4749 of *LNCS*, pages 193–206, Vienna, Austria, 2007.
- [11] K. Vengataraghavan and R. Sundarraj. On Integrating Time Preferences and Concession Models into Cloud Computing Negotiations. In *Intl. Symp. on Cloud and Services Computing*, pages 157–163, Mangalore, India, 2012.