

Cloud Computing con herramientas libres para evaluación de modelos de despliegue híbrido

Nelson Rodríguez¹, Adriana Valenzuela², María Murazzo³, Susana Chávez⁴,
Adriana Martín⁵, Daniela Villafañe⁶, Facundo Gonzalez⁷

Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.

Complejo Islas Malvinas. Cereceto y Meglioli. 5400. Rivadavia. San Juan

¹nelson@iinfo.unsj.edu.ar ²franciscaadriana.valenzuela@gmail.com ³marite@unsj-cuim.edu.ar

⁴schavez@iinfo.unsj.edu.ar ⁵arianamartinsj@gmail.com ⁶villafane.unsj@hotmail.com

⁷facu_jgg@hotmail.com

Resumen

El avance notable de tecnologías como la computación distribuida, Internet y grid computing, han posibilitado que Cloud Computing forme parte de un nuevo modelo de computación y de negocios. Cloud Computing está transformando los modos tradicionales de cómo las empresas utilizan y adquieren los recursos de Information Technology (IT). Representa un nuevo tipo del valor de la computación en red. Entrega mayor eficiencia, escalabilidad masiva y más rápido y fácil desarrollo de software. Los nuevos modelos de programación y la nueva infraestructura de IT están impulsando nuevos modelos de negocios. Luego de un auge inicial de los Cloud Públicos las empresas han comenzado a montar Cloud híbridos que ofrecen las ventajas de Cloud Computing sumado a la privacidad de los datos que consideren estratégicos. En la actualidad, datos de sistemas propios están almacenados en servidores privados y muchos otros como sitios Web y servicios de e-mails se encuentran en algún proveedor remoto. Una solución de Cloud híbrido permite la integración de ambos sistemas.

No obstante las ventajas expuestas, no se han realizado estudios profundos de escalabilidad, eficiencia e interoperabilidad. Tampoco se han definido parámetros para este tipo de

soluciones, que permitan a los desarrolladores proponer la arquitectura más conveniente.

Palabras clave: *Cloud Computing, Hybrid Cloud Computing, Open Source, IaaS, SaaS, PaaS*

Contexto

El presente trabajo se encuadra dentro del área de I/D Innovación en Sistemas de Software, y se enmarca dentro del proyecto de investigación Implantación de un ambiente de Cloud Computing para integración de recursos, presentado en diciembre de 2013 y que tiene como unidades ejecutoras al Departamento e Instituto de Informática de la FCEFYN de la UNSJ. Esta propuesta se ha presentado en una nueva convocatoria, pero los resultados de evaluación la misma se espera que estén para abril del presente año.

Introducción

En los últimos años las IT han sufrido cambios drásticos y acelerados, se pasó de un modelo centralizado, a un modelo ampliamente distribuido en donde conviven data centers, computadoras personales, clusters, dispositivos móviles, etc. Estos cambios responden al acelerado flujo de datos y operaciones que surgen al interior de cada organización. Como

consecuencia de ello, Cloud Computing (CC) apareció como un modelo tendiente a proveer una serie de servicios que utilizan eficientemente estos recursos.

Las formas tradicionales en que las empresas utilizan y adquieren los recursos tecnológicos han evolucionado con el auge de CC, ofreciendo mayor eficiencia, escalabilidad masiva y más rápido y fácil desarrollo de software [1].

CC representa un modelo de prestación de servicios de negocio y tecnología que permite a cualquier usuario acceder a un catálogo de servicios estandarizados y responder a las necesidades de su negocio de forma flexible y adaptativa, en caso de demandas no previsibles o de picos de trabajo, pagando únicamente por el consumo efectuado.

CC no es un desarrollo revolucionario reciente, sino el resultado de la evolución de varias tecnologías tales como utility computing, computación bajo demanda, computación elástica y grid computing [2].

La característica básica de este modelo es que los recursos y servicios informáticos, tales como infraestructura, plataforma y aplicaciones, son ofrecidos y consumidos como servicios a través de la Internet sin que los usuarios tengan que tener ningún conocimiento de lo que sucede detrás [3].

El Instituto Nacional de Estándares y Tecnología de EE.UU (NIST) define CC como un modelo que habilita acceso a red ubicuo, conveniente, bajo demanda para compartir un conjunto de recursos configurable, que pueden ser rápidamente provistos y liberados con mínimo esfuerzo o interacción del proveedor de servicios. Además, se resaltan los cinco (5) aspectos claves de CC, a saber: auto servicio bajo demanda, acceso a la red ubicua, un conjunto de recursos independiente de la ubicación, rápida elasticidad y servicio a medida [4].

CC es un esquema del tipo XaaS o Everything as a Service. Se suele dividir a CC en las siguientes capas: Software como Servicio (SaaS), Plataforma como Servicio (PaaS) e Infraestructura como Servicio (IaaS) [5], aunque suelen aparecer otras capas como Hardware como Servicio (HaaS) y Database como Servicio (DaaS), entre otras.

Frank Gens, Analista en IDC dice: “Un reciente estudio entre Ejecutivos de IT, CIOs y colegas en las líneas de negocio muestra que CC está entrando en un período de amplia adopción. La crisis económica amplificará la adopción de Cloud. Este modelo ofrece una manera más barata para que el negocio use y adquiera tecnología. Esta ventaja es importante para los pequeños y medianos negocios, un sector que será clave en cualquier plan de recuperación” [6].

Algunas problemáticas que surgen sobre este modelo apuntan a la confidencialidad de los datos y otras cuestiones que en general son inherentes a qué puede ser público y qué debe ser privado. Al respecto, un Cloud privado puede configurar la confidencialidad de los datos en función de las necesidades de los usuarios, pero la escalabilidad de este Cloud puede verse limitada. Esta limitación tiene que ver con la forma en que se suministran los recursos. Si los recursos del Cloud están ocupados en el cumplimiento de solicitudes, una nueva solicitud no puede efectuarse.

Para evitar demoras se puede trasladar el manejo de peticiones de baja prioridad al Cloud público si el privado no tiene disponibilidad. Surge así una nueva alternativa a través de un Cloud híbrido. NIST define una infraestructura de Cloud híbrido como: "Una composición de dos o más infraestructuras distintas de Clouds (privado, comunitario o público) que siendo entidades únicas están unidas por tecnología estandarizada o propietaria que

permite la portabilidad de datos y aplicaciones" [4].

Este tipo de computación se utiliza cuando una organización tiene períodos de estacionamiento de su actividad. Esto es, una vez que la infraestructura tecnológica interna no puede hacer frente a los desafíos actuales, algunas facilidades son transferidas a un Cloud público (por ejemplo, grandes cantidades de información estadística, que en su forma cruda no tiene un gran valor para la empresa) permitiendo proporcionar al usuario acceso a los recursos empresariales (para el Cloud privado) a través del Cloud público [7].

Una solución híbrida permitirá delegar en el Cloud público solo cierta parte de la información, de forma tal de que no represente una amenaza si ésta es recuperada por usuarios externos y al mismo tiempo permitirá una alta disponibilidad a la información almacenada al interior de la empresa y una disminución en el consumo de recursos [8].

Sin embargo, las ventajas presentadas no son suficientes para lograr una estrategia gestión eficaz para la implementación de Cloud híbrido; existen muchísimos problemas a resolver.

A pesar de que el Cloud público y el privado funcionan “emparejados”, pueden existir diferencias de diseño. Cuantos mayores sean estas diferencias, más difícil será gestionar múltiples clouds como una sola entidad.

El Cloud híbrido favorece la redundancia ahorrando costos y mitigando los riesgos de utilizar múltiples centros, pero es necesario resolver la sincronización y el manejo de versiones.

Mantener y demostrar el cumplimiento de lo pactado es más difícil en un Cloud híbrido. No solo porque hay que asegurarse de que el proveedor público y el privado estén cumpliendo, sino que

además hay que cerciorarse que los medios de coordinación entre los dos Clouds sean compatibles. Por ejemplo si se trabaja con medios de pagos electrónicos ambos Clouds deben soportar dicho sistema de pago.

Con el proveedor de Cloud público se realiza un acuerdo de servicio (SLA), que tiene valores pautados. Para eso se deben recopilar datos de trabajo del Cloud privado bajo carga de trabajo realista para estimar las necesidades del Cloud público, y además se deben buscar los problemas con la integración del Cloud público y privado que podrían interrumpir el servicio.

En cuanto a la seguridad, CC híbrido en particular, utiliza nuevas APIs, demanda configuraciones complejas, y requiere nuevos conocimientos y habilidades de los administradores de sistema. Estos factores introducen nuevos tipos de amenazas. Esto es así pues los Clouds híbridos son un sistema complejo, sobre los que los administradores tienen una limitada experiencia y esto crea riesgo.

Los controles de seguridad tales como la autenticación, la autorización y la gestión de identidad, se deberán trabajar tanto en el sector privado como en el público, integrando estos protocolos de seguridad, ya sea replicando los controles en los dos Clouds y manteniendo los datos de seguridad sincronizados, o utilizando un servicio de gestión de identidad para ambos Clouds.

El soporte de red resulta clave en todas las actualizaciones y se debe revisar los parámetros de latencia, y velocidad de transferencia efectiva si se quiere que la arquitectura de Cloud híbrido ofrezca las ventajas que se publicitan.

La implementación de un Cloud híbrido presenta más que sólo desafíos técnicos, los administradores de IT también tienen que hacer frente a una variedad de problemas. Al entender y vencer los

obstáculos mencionados, el Cloud híbrido podría ofrecer más recompensa que el riesgo de implementarlo.

En ambientes académicos existen pocos proyectos sobre Cloud híbridos, a pesar de ello, algunas publicaciones científicas pueden encontrarse, enfocadas a desarrollos específicos para resolver problemas puntuales, pero que no realizan un análisis exhaustivo de las tecnologías. Existen a su vez propuestas de empresas, sobre la filosofía de trabajo en entornos híbridos y algunas proyecciones pero en ningún caso implican trabajos de investigación profundos.

Los gestores de infraestructura no comparten una arquitectura común y la interoperabilidad no está garantizada. Además hacen falta estudios sobre el comportamiento de estos gestores en función de las necesidades del negocio en cuanto a la escalabilidad, eficiencia e interoperabilidad, que es el objetivo del proyecto presentado por el grupo de investigación.

Líneas de Investigación, Desarrollo e Innovación

Cloud Computing involucra una gran variedad de tecnologías, además el modelo de despliegue híbrido incluye al modelo público y al privado. Esto determina que los investigadores del grupo tengan que resolver múltiples problemas.

La complejidad del enfoque híbrido radica en la forma de planificar la asignación de procesos y datos entre ambas infraestructuras, gestión de seguridad y otros aspectos, que deben ser solucionados para obtener un balanceo de carga óptimo.

El grupo de investigación debe realizar el análisis de la escalabilidad, eficiencia,

interoperabilidad de la tecnología de CC híbrido utilizando herramientas libres, construyendo nuevas herramientas que permitan evaluar los aspectos mencionados.

Problemas a resolver en un futuro

Además de estudiar los problemas antes mencionados sobre la gestión de configuración, el control de cambios y la seguridad específica para Cloud híbrido, se plantea para un futuro abordar temáticas vinculadas a la gestión de fallos y la gestión presupuestaria, reusabilidad, gestión de réplicas, extensibilidad, entre otras; la mayoría de éstos son desafíos previstos para futuros proyectos.

Resultados y Objetivos

Resultados Obtenidos

A pesar de que esta línea de investigación recién ha sido presentada para los años 2014 y 2015, el grupo ha realizado publicaciones en el área durante el último año, seis (6) trabajos de investigación en diferentes Congresos y Jornadas: dos trabajos en el Workshop de Investigadores en Ciencias de la Computación 2013 [9] [10], dos (2) trabajos en el Congreso Argentino de Ciencias de la Computación 2013 [2] [11], dos (2) trabajos en las 1eras Jornadas de Cloud Computing [1] [12], además se realizaron dos (2) publicaciones en revistas [13] [14]. También se han aprobado dos tesis de grado, una sobre Google AppEngine y otra sobre Mobile Cloud Computing.

Objetivos

El objetivo del grupo de investigación es realizar el análisis de la escalabilidad, eficiencia, interoperabilidad de la tecnología de Cloud Computing híbrido utilizando herramientas libres.

Formación de Recursos Humanos

El equipo de trabajo está compuesto por los siete (7) docentes-investigadores que figuran en este trabajo.

Durante 2013 se aprobaron dos (2) tesinas de grado. Además se están realizando dos (2) tesinas de licenciatura una sobre Mobile Cloud Computing y otra sobre Hybrid Cloud Computing. Se espera realizar también dos (2) tesis de maestría una sobre Metodologías de desarrollo aplicadas a SaaS y otras sobre bases de datos NoSQL y aumentar el número de publicaciones. Por otro lado también se prevé la divulgación de varios temas investigados por medio de cursos de postgrado y actualización o publicaciones de divulgación.

Referencias

- [1] Rodríguez, Murazzo, Chávez, Valenzuela, Martín, Villafañe. Aspectos claves para el desarrollo de aplicaciones para Mobile Cloud Computing I JCC. UNLP. Jun. 2013. La Plata- Bs As.
- [2] Rodríguez, Murazzo, Villafañe, Valenzuela, Martín, Chávez. Una propuesta para la incorporación de Cloud Computing en la currícula de Grado. XiX CACIC. Oct. 2013. Mar del Plata. Bs As
- [3] Weiss, Computing in the clouds. netWorker Vol 11, Issue 4, page. 16-25. Dec 2007.
- [4] NIST Special Publication 800-145. The NIST Definition of Cloud Computing. <http://csre.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>
- [5] S. Rao, N. Rao, Kumari. Cloud Computing: An Overview. Journal of Theoretical and Applied Information Technology. Vol 9 N1. 2009.
- [6] IDC, "IDC Finds Cloud Computing". "Entering Period of Accelerating

Adoption and Poised to Capture IT Spending Growth Over the Next Five Years". <http://www.idc.com/research/cloudcomputing>.

- [7] Parthipan, Sriprasadh, Maheshkumar. Secure Information Transaction In Hybrid Cloud Computing. Information Communication and Embedded Systems (ICICES). 2013.
- [8] Hernández Ramírez. Sistema de almacenamiento de archivos con tolerancia a fallos utilizando Cloud Híbrido. Tesis de Maestría en Computación. Victoria, Tamaulipas, México. 2011.
- [9] Rodríguez, Murazzo, Villafañe, Alves, Medel. Integración de Computación Heterogénea con Hadoop para Cloud Computing. XV WICC. Abril 2013. Paraná. Entre Ríos.
- [10] Martín, Chávez, Rodríguez, Valenzuela, Murazzo Bases de Datos NoSql en Cloud Computing. XV WICC. Abril 2013. Paraná. Entre Ríos.
- [11] Murazzo, Rodríguez, Villafañe, Gallardo. Desarrollo de aplicaciones colaborativas para Cloud Computing. XIX CACIC. Oct. 2013. Mar del Plata. Bs As.
- [12] Murazzo, Rodríguez, Villafañe, González. Perspectivas en el análisis de grandes volúmenes de datos en el Cloud. I JCC. UNLP. Jun. 2013. La Plata- Bs As.
- [13] Rodríguez, Murazzo, Chávez, Valenzuela, Martín, Villafañe. Key aspects for the development of applications for Mobile Cloud Computing. Journal of Computer Science & Technology. Vol. 13 - No. 3 – Dec. 2013. Special Issue on "I JCC 2013".
- [14] Murazzo, Rodríguez, Villafañe, González. Perspectives in processing large amounts of information using Cloud. Journal of Computer Science & Technology. Vol. 13 - No. 3 – Dec. 2013. - ISSN 1666-6038. Special Issue on "I JCC 2013".