

Performance de cloud computing para HPC: despliegue y seguridad.

Brian Galarza¹, Gonzalo Zaccardi^{1,2}, Maximiliano Belizán¹, Jorge Osio¹, David Duarte¹,
Martin Morales^{1,3}, Diego Encinas^{1,4}

¹Instituto de Ingeniería y Agronomía - Universidad Nacional Arturo Jauretche

²Dirección Tecnológica – Universidad de Palermo

³Unidad CodApli - Facultad Regional La Plata - UTN

⁴Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática - UNLP – Centro Asociado CIC

bgalarza@unaj.edu.ar, gzacca@palermo.edu, maximiliano.h.belizan@gmail.com,
josio@unaj.edu.ar, davito.duarte.22@gmail.com, martin.morales@unaj.edu.ar,
dencinas@unaj.edu.ar

Resumen

El objetivo de esta línea de investigación es el estudio de la performance de las arquitecturas tipo cloud a través del despliegue de IaaS y utilización de IaaS públicos, en particular en el área de cómputo paralelo de altas prestaciones (HPC). Enfocando a la obtención de herramientas que permitan predecir la eficiencia del sistema ante posibles escenarios. Analizando los diferentes componentes del sistema que pueden influir en las prestaciones significativamente y pueden llegar a modelarse y/o configurarse.

Palabras clave: *Arquitecturas Multiprocesador. Cloud Computing. OpenStack. Simulación.*

Contexto

Se presenta una línea de Investigación que es parte del Proyecto de Investigación

“Modelado y Simulación en Cómputo de Altas Prestaciones (HPC). Aplicaciones en arquitecturas multiprocesador, sistemas paralelos y redes de datos” de la Universidad Nacional Arturo Jauretche (UNAJ), acreditado por resolución interna 186/15. Además, el Proyecto integra el Programa “Tecnologías de la información y la comunicación (TIC) en aplicaciones de interés social” de la UNAJ.

Por otra parte, se tiene financiamiento en el marco del programa “Universidad, Diseño y Desarrollo Productivo” del Ministerio de Educación a través de varios proyectos aprobados en la UNAJ.

Introducción

Cloud Computing es un paradigma que está en constante crecimiento durante estos últimos años, cada vez más compañías y grupos de investigación trabajan en conjunto con el fin de explotar las oportunidades ofrecidas por el mismo

[1]. Dicho paradigma ofrece muchas ventajas, tales como el bajo costo de implementación, ya que no se necesitan computadoras de última tecnología debido a que éstas trabajan conjuntamente (Clustering) con la posibilidad de escalar horizontalmente de manera sencilla. Además, hay software Open Source disponible para los nodos en el clúster como las infraestructuras Eucalyptus, OpenNebula, CloudStack u OpenStack integradas con GNU/Linux y compatibles, por ejemplo, con Amazon WebServices.

Despliegue de IaaS

Las comunicaciones en Cloud Computing son una parte fundamental del paradigma que consisten en utilizar distintos nodos y lograr hacerlos funcionar conjuntamente.

Para lograr una comunicación entre estos nodos se propone utilizar OpenStack [2].

OpenStack es un software de código abierto que permite la implementación de, por ejemplo, una “Infraestructure as a Service” (IaaS) a través de múltiples servicios que, de manera coordinada, cumplen diferentes propósitos para lograr el correcto funcionamiento de dicha infraestructura. Algunos de los servicios ofrecidos por OpenStack son: hypervisor (Nova), autenticación (Keystone), Imágenes (Glance), Dashboard (Horizon) y block storage (Cinder). Según las necesidades se pueden requerir de ciertos servicios u otros. La Arquitectura básicamente consiste en dos tipos de nodos: “Compute Node” y “Controller Node”. Se llaman Compute Node a todos aquellos que se encargan del procesamiento de servicios específicos mientras que Controller Node es aquel que comunica a cada uno de los anteriores [3] [4] [5].

La implementación de estas infraestructuras ofrece ventajas en las cuales los clústers virtualizados trabajan en conjunto ofreciendo un buen rendimiento a bajos costos y con posibilidad de escalabilidad al poder agregar mayor cantidad de nodos para procesamiento de manera sencilla.

Fuel es una herramienta Desarrollada por Mirantis en la cual se ejecuta un script que permite configurar, de manera más amigable respecto a OpenStack, los recursos que se desean otorgar a la infraestructura, como la cantidad de nodos, los núcleos de procesador, la memoria RAM, entre otros [6].

Fuel trabaja con un nodo master el cual es el encargado de controlar a los nodos slaves que contendrán la infraestructura OpenStack. Es decir, desde el nodo Fuel Master se indican qué paquetes se van a instalar en cada nodo slave (Glance, Nova-Compute, Keystone, etc.) para luego en los slaves tener armados los nodos compute y controller, sin necesidad de realizar configuraciones manuales en cada uno de los mismos.

Seguridad en arquitecturas de Cloud Computing

Aunque Cloud Computing tiene muchos beneficios, el uso inadecuado puede conllevar a pérdidas de datos sensibles para el usuario y el proveedor. Entre los principales riesgos y problemas de seguridad/privacidad se puede destacar los siguientes: Falta de control de datos, Ambigüedad de responsabilidad entre el usuario y el proveedor, autenticación y autorización, error de aislamiento, cumplimiento y riesgos legales, manejo de incidentes de seguridad, vulnerabilidad de la interfaz de administración, protección de aplicaciones y de datos, indisponibilidad del servicio, bloqueo del proveedor, eliminación de los datos

inseguros o incompletos, visibilidad y auditoría y seguridad [7] con la virtualización (ataques entre MVs, MVs inactivas desactualizadas, falta de monitoreo y registro de MVs, Host con MVs seguras e inseguras y ataques DoS por uso no autorizado de recursos compartidos) [8].

Es importante crear un modelo de seguridad de datos, para prevenir los problemas antes mencionados, que proporcionen seguridad en los: datos para procesar (encriptación homomórfica, ejemplo: Unpadded RSA), datos de transmisión (certificados SSL, ejemplo: HTTPS), datos de almacenamiento (encriptación, ejemplo: AES), administración de claves (ejemplo: HSM), Autenticación y autorización (autenticación de dos factores, ejemplo: Token), protección de usuario (alentar el encriptado de datos sensibles antes de subir) y controles de uso (Firewalls) [9]. OpenStack cuenta con un servicio llamado Keystone que ayuda en la autenticación y autorización distribuida de clientes, esto con apoyo de la encriptación de la base de datos, comunicación sobre capas seguras y el uso de Firewall, como Iptables, puede ayudar a proveer un Cloud más confiable [10].

Finalmente se propone la implementación de seguridad en la capa física. Ésta, va más allá del uso de OpenStack, ya que se busca cierta restricción al acceso físico de los equipos proveedores del servicio Cloud de OpenStack.

Estas implementaciones, más las actualizaciones en los distintos dispositivos (equipos, router, etc.) y software (OpenStack, iptables, etc) que se utilicen.

Líneas de Investigación, Desarrollo e Innovación

Temas de Estudio e Investigación

- Arquitecturas multiprocesador para procesamiento paralelo: multiprocesador de memoria compartida, multiprocesador on-chip de memoria distribuida. Multicore, Clusters, Clusters de multicore. Grid. Cloud.
- Plataformas de software para implementar y administrar Clouds públicos, privados e híbridos.
- Seguridad y privacidad en Cloud Computing.

Resultados y Objetivos

Investigación experimental

- Implementación de un IaaS encargado de realizar operaciones en procesamiento paralelo aumentando la eficiencia y reduciendo los costes generados.
- Implementación de OpenStack Dashboard y de un sistema desarrollado para poder controlar/administrar de manera visual (web) y más básica cada uno de los servicios.
- Utilización de Fuel para administrar OpenStack como sistema de administración de nube (Cloud Computing) a partir de la infraestructura de 2 nodos compute y 1 controller [11].

- Lograr escalabilidad agregando nuevos nodos compute a la infraestructura obteniendo un mayor performance en el sistema.
- Analisis del rendimiento de un Cloud privado en la ejecución de instancias personalizadas.
- Ejecutar sobre el modelo de nodos implementado en OpenStack la distribución GNU/Linux Hetnux, desarrollada en la UNAJ en el marco del programa “Universidad, Diseño y Desarrollo Productivo 2014”.
- Ejecutar diferentes benchmarks en la infraestructura desplegada sobre OpenStack para así comparar los resultados obtenidos con infraestructuras similares montadas en Azure y Amazon y poder realizar un análisis del rendimiento en cada caso.
- Implementación de un modelo de seguridad en un IaaS.
- Comprobar las mejoras prácticas para la seguridad de Cloud y modificar el modelo de seguridad, de ser necesario, para una mayor robustez.
- Como resultado del modelo de seguridad se espera un servicio de infraestructura Cloud que proteja los servicios usados y los datos generados por los usuarios.

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D se participa en el dictado de la carrera de Ingeniería Informática de la UNAJ. También aportan trabajos de alumnos de las materias Sistemas Operativos 1, Redes de Computadoras 2, Programación en

Tiempo Real y Organización y Arquitecturas de Computadoras.

Durante 2017 se han realizado 3 publicaciones nacionales.

En esta línea de I/D existe cooperación a nivel nacional. Hay 2 investigadores realizando carreras de postgrado, 2 becarios de grado (EVC CIN y entrenamiento CIC) y 2 alumnos avanzados de grado colaborando en las tareas.

Referencias

1. Kondo, D., Javadi, B., Malecot, P., Cappello, F., Anderson, D. P.: “Cost-benefit analysis of Cloud Computing versus desktop grids”. In: IPDPS '09 Proceedings. IEEE International Symposium on Parallel and Distributed Processing, Washington, USA (2009).
2. OpenStack Cloud Software: Open source software for building private and public clouds. <http://www.openstack.org>. Febrero 2015.
3. Galarza, B.; Tuamá, C.; Zaccardi, G.; Encinas, D.; Morales, M. “Implementaciones de Cloud Computing y aplicaciones en el ámbito universitario”. I Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2013). Ciudad de Córdoba, Argentina.
4. Zaccardi, G.; Galarza, B.; Encinas, D.; Morales, M. “Implementación de Cloud Computing utilizando OpenStack”. II Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2014). Ciudad de San Luis, Argentina.
5. Galarza, B.; Zaccardi, G.; Encinas, D.; Morales, M. “Análisis de despliegue de una IaaS utilizando Openstack”. XXI Congreso Argentino de Ciencias de la Computación (CACIC 2015). Ciudad de Junín, Argentina.
6. OpenStack Deployment Fuel. <https://www.mirantis.com/products/miranti>

- s-openstack-software/openstack-deployment-fuel/. Febrero 2016
7. Cloud Standards Consumer Council <http://www.cloud-council.org/deliverables/CSCC-Security-for-Cloud-Computing-10-Steps-to-Ensure-Success.pdf>. Marzo 2017
 8. Mishra, A.; Mathur, R.; Jain, S.; Rathore, J. “Cloud Computing Security”. *International Journal on Recent and Innovation Trends in Computing and Communication* ISSN 2321–8169 Volume: 1 Issue: 1 36 – 39. 2013
 9. Mohamed, E.; Abdelkader, H.; El-Etriby, S. “Data Security Model for Cloud Computing”. *Journal of Communication and Compute*. ISSN: 1047-1062. 2013
 10. Balu, V.; Mary, L. “A Model of Security Architecture on Private Cloud Using OpenStack”. *International Journal on Recent and Innovation Trends in Computing and Communication* ISSN: 2321-8169 Volume: 3 Issue: 2 587-590. 2015.
 11. G. Zaccardi, B. Galarza, M. Morales, D. Encinas. Despliegue y ejecución de un cloud privado. IV Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2016). Ciudad de Salta, Argentina. ISSN 2347-0372. 2016