

ASPECTOS DE INGENIERÍA DE SOFTWARE, BASES DE DATOS RELACIONALES, Y BASES DE DATOS NO RELACIONALES Y BASES DE DATOS COMO SERVICIOS EN LA NUBE PARA EL DESARROLLO DE SOFTWARE HÍBRIDO.

Luciano Marrero , Pablo Thomas , Ariel Pasini , Rodolfo Bertone , Eduardo Ibáñez ,
Verónica Aguirre , Marisa Panizzi¹, Verena Olsowy , Fernando Tesone , Patricia Pesado 

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 La Plata Buenos Aires
Centro Asociado CIC
¹ Universidad de Morón

*{lmarrero, pthomas, apasini, pbertone, eibanez, vaguirre, volsowy, ftesone,
ppesado}@lidi.info.unlp.edu.ar,*
¹marisapanizzi@outlook.com

RESUMEN

Se presenta una línea de investigación que tiene por objeto estudiar las problemáticas actuales que afronta los procesos de Ingeniería de Software y Bases de Datos ante una gran variedad aplicaciones multiplataformas (Web y móvil), que son de uso constante por millones de usuario simultáneamente y de cualquier punto geográfico. En este contexto, las metodologías ágiles, son sin duda, las más utilizadas en el proceso de desarrollo del Software actual. El modelo {relacional de Bases de Datos (Codd 1970) [5], es el modelo predominante de almacenamiento de información. Sin embargo, la idea de considerar que un único modelo de datos pueda adaptarse de forma eficiente a todos los requerimientos, ha sido discutida. Surgen así, otros motores de Bases de Datos que poseen implementaciones propias no relacionales y se denominan Bases de Datos NoSQL (No solo SQL). Estas Bases de Datos son un

complemento y/o alternativa a las Bases de Datos Relacionales en el proceso de diseño de la información. En la actualidad, los proveedores de Bases de Datos No Relacionales ofrecen implementaciones alternativas en la nube. Con esto se cuenta con un conjunto de herramientas importantes para el desarrollo de aplicaciones móviles. Esto implica, tener en cuenta aspectos que hace algunos años no eran considerados, tales como, movilidad, geolocalización, generación de grandes volúmenes de información y la diversidad de los dispositivos electrónicos involucrados [1, 2, 9, 10].

Palabras claves: Metodologías ágiles, Bases de Datos Relacionales, Bases de Datos NoSQL, Bases de Datos en la nube, Aplicaciones Móviles.

CONTEXTO

Esta línea de Investigación forma parte del proyecto (2018-2021) “Metodologías, técnicas y herramientas de ingeniería de software en

escenarios híbridos. Mejora de proceso.”, en particular del subproyecto “Ingeniería de Software para escenarios híbridos”, del Instituto de Investigación en Informática LIDI (III-LIDI) de la Facultad de Informática UNLP, acreditado por el Ministerio de Educación de la Nación.

Hay cooperación con Universidades de Argentina y se está trabajando con Universidades de Europa en proyectos financiados por el Ministerio de Ciencia y Tecnología de España y la AECID.

Se utilizan los recursos de Hardware y Software disponibles en el III-LIDI para diseñar, desarrollar y probar diferentes soluciones a problemáticas relacionadas con escenarios a investigar. Como resultado de esto, se espera obtener métricas reales que sirvan como referencia para los investigadores en la comparación de resultados.

Las publicaciones científicas generadas y la transferencia continua de resultados concretos, validan esta línea de investigación.

1. INTRODUCCIÓN

“Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso”, 2018-2021 del Programa de Incentivos, es un proyecto que propone profundizar las investigaciones que se vienen realizando en el III-LIDI y extender la mirada a nuevos desafíos y cambios que están en gestación. Se organiza en tres subproyectos que permiten atender de manera ordenada el objetivo general propuesto:

“SP1 - Ingeniería de Software para escenarios híbridos”, SP2 - Gobernanza Digital. Mejora de Procesos.” Y “SP3 - Metodologías y

herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos [19].

Este artículo se centra en el subproyecto “SP1 - Ingeniería de Software para escenarios híbridos”. Se orienta a la investigación de metodologías y técnicas de la Ingeniería de Software, Bases de Datos Relacionales, Bases de Datos NoSQL y Bases de Datos en la nube, con énfasis en los escenarios híbridos y las nuevas aplicaciones Web móvil para cualquier dispositivo informático.

La disponibilidad de la información en todo momento (internet), las nuevas tecnologías móviles y el desarrollo de aplicaciones híbridas, han cambiado las necesidades en el mundo informático. Se presentan así, un conjunto de nuevas alternativas y desafíos para la Ingeniería de Software, lo que conlleva al desafío de establecer nuevos procesos de diseño y nuevas arquitecturas de Hardware, cambiando así, la forma de pensar y almacenar los datos. Las Bases de Datos NoSQL representan una alternativa en la evolución del almacenamiento de datos, complementándose con una generación de tecnología móvil Web que debe responder eficientemente a las exigencias del usuario [9, 10, 11 y 12].

Cada vez, son más los recursos o servicios informáticos que poseen variantes en la nube (Cloud Computing). Esta tecnología, permite el ahorro en la administración de servicios y equipos. Es una abstracción de los recursos tecnológicos en donde se puede acceder a la información independientemente del punto geográfico en el que se encuentre el dispositivo. [4, 5, 8, 15 y 16].

En la actualidad, no solo se debe responder de manera óptima al número de usuarios finales, sino que también, cada usuario ha aumentado

sus exigencias al momento de utilizar y/o catalogar a una aplicación Web o móvil Web. Esto supone que la escalabilidad de los recursos y el rendimiento se han convertido en auténticos retos para la Ingeniería de Software y los proveedores de Bases de Datos.

En este contexto, las tecnologías de Bases de Datos No Relacionales (NoSQL), son un complemento o alternativa para satisfacer las demandas actuales. Las bases de datos NoSQL, surgen como respuesta para la administración de grandes volúmenes de información. Son altamente escalables y no respetan estrictamente las propiedades ACID (Atomicidad, Consistencia, Aislamiento y Durabilidad) [4 y 5]. NoSQL propone un sistema llamado BASE (Básicamente Disponible, Estado Suave y Consistencia Eventual) [10, 13 y 14]. Existe una gran variedad de motores de Bases de Datos NoSQL que en general, se pueden catalogar en cuatro categorías de almacenamiento no estructurado de información: Clave-Valor, Documental, Orientado a Columnas y Orientado a Grafos. Herramientas como MongoDB (documental), Apache Cassandra (orientada a columnas), Redis (clave-valor) y Neo4j (grafos) son ejemplos de esta nueva generación de motores de Bases de Datos [9, 12 y 18].

Además, existen proveedores que brindan alternativas de Bases de Datos NoSQL en la nube.

Cloud Firestore, es un motor de Base de Datos en la nube con almacenamiento documental que pertenece a un conjunto de servicios que brinda Google para el desarrollo de aplicaciones Web y móvil.

MongoDB Atlas, es una Base de Datos en la nube con almacenamiento documental. Es proporcionada por MongoDB, es decir, que se

encuentra operada y mantenida por MongoDB. Este servicio incluye de forma automática, la configuración de los servidores y de todo el entorno necesario para la Base de Datos. Además, es compatible con diferentes proveedores de servicios en la nube, como son: Amazon Web Services (AWS), Google Cloud Platform (GPC) y Microsoft Azure.

DataStax (Astra), es otro ejemplo de un servicio de Bases de Datos en la nube, en este caso impulsado por Apache Cassandra [7, 11, 13 y 17].

Como parte de las investigaciones realizadas, el análisis del proceso de despliegue del Software es otro aspecto de estudio. Las PyMES en Argentina representan casi el 80% de la industria del software y dada la necesidad de ser competitivas deben mejorar sus métodos y procesos de trabajo. En la mayoría de las empresas el proceso de despliegue no se realiza de manera sistemática y controlada, esto impacta en la finalización del proyecto y la no aceptación del producto final, ocasionando inconvenientes que generan rehacer el trabajo y baja de productividad en su proceso. Ante estas dificultades, se propone realizar un modelo de proceso de despliegue de sistemas de software que permita a las PyMES mejorar la ejecución del proceso de despliegue.

La comunicación y la sincronización del trabajo continúa siendo un pilar fundamental para el éxito de un proyecto. La utilización de repositorios de información, por ejemplo, GIT, permiten realizar un control de versiones distribuido, trabajando en modo offline o en modo online, con la facilidad de disponer herramientas específicas para la resolución de conflictos entre versiones [3, 6 y 16].

Todas las particularidades previamente mencionadas conducen a nuevas prácticas y metodologías que promueven el crecimiento de la Ingeniería de Software y la gestión de Bases de Datos como disciplina.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

- Metodologías y Técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para escenarios híbridos.
- Investigar los distintos tipos de almacenamiento no estructurado de información (documental, orientado a columnas, clave-valor y grafos).
- Investigar motores de Bases de Datos No Relacionales (NoSQL). MongoDB, Apache Cassandra, Redis, Neo4j, entre otros.
- Investigar Bases de Datos en la nube. Cloud Firestore (Google), MongoDB Atlas (MongoDB), DataStax Astra (Apache Cassandra), entre otras.
- Investigar nuevos conceptos de tipos de Bases de Datos, como, por ejemplo: NewSQL y Bases de Datos de Serie Temporales.
- Desarrollo de casos de estudio, pruebas de comparación y rendimiento de Bases de Datos Relacionales, No Relacionales y Bases de Datos en la nube.
- Modelo de procesos para el despliegue / puesta en producción de sistemas de software
- Repositorios GIT.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados esperados/obtenidos se pueden resumir en:

- Capacitación continua de los miembros de las líneas de investigación.
- Estudio y análisis de Bases de Datos no relacionales.
- Estudio y análisis de Bases de Datos en la nube.
- Analizar, comparar y determinar escenarios para los distintos tipos de almacenamiento no estructurado de información.
- Comparar y analizar resultados para diversos casos de estudio entre Bases de Datos relacionales, Bases de Datos NoSQL y Bases de Datos en la nube.
- Definición de procesos de Gestión de Incidencias utilizando repositorios GIT.
- Análisis de metodologías para la interoperabilidad de sistemas web y aplicaciones móviles.

Algunas de las transferencias realizadas por el III-LIDI relacionadas con este proyecto, entre otras, son:

- Aplicación móvil para la comunidad de la Facultad de Informática de la UNLP con información sobre horarios de finales, planes de estudio, calendario académico, las clases en tiempo real y las últimas novedades.
- Diseño de Bases de Datos para Instituciones Provinciales.
- Diseño y Gestión de Sistemas de congresos, utilizado por la RedUNCI y por otras entidades.

- Sistemas de Gestión Administrativa de Instituciones Universitarias (Sistema de inscripción y seguimiento de alumnos de la Facultad de Informática, Sistema de gestión administrativa de asignaturas, entre otros).

4. FORMACIÓN DE RECURSOS HUMANOS

Los integrantes de esta línea de investigación dirigen Tesinas de Grado y Tesis de Postgrado en la Facultad de Informática, y Becarios III-LIDI en temas relacionados con el proyecto. Además, participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP.

5. BIBLIOGRAFÍA

1. Ingeniería de Software. Un Enfoque Práctico. Séptima Edición. Roger S. Pressman. McGraw-Hill Interamericana Editores, S.A. 2010. ISBN: 978-607-15-0314-5
2. Ingeniería de Software Teoría y Práctica. Pflieger Shari Lawrence. Pearson / Prentice Hall. 2002. ISBN: 9789879460719
3. Ingeniería de Software. Novena Edición. Ian Sommerville. Addison Wesley / Pearson. 2011. ISBN: 978-607-32-0603-7
4. Ingeniería de Software Clásica y Orientada a Objetos. Sexta Edición. Stephen R. Schach. Mc Graw Hill Interamericana Editores S.A. 2006. ISBN: 970-10-5636-1.
5. Diseño Conceptual de Bases de Datos, un enfoque de entidades-interrelaciones. Carlo Batini, Stefano Ceri, Shamkant B. Navathe.. Addison-Wesley / Díaz de Santos. ISBN 0-201-60120-6 (1994).
6. Administración de Proyectos. Guía para el Aprendizaje. Francisco Rivera Martínez, Gisel Hernández Chávez. Prentice Hall / Pearson. 2010. ISBN: 978-607-442-620-5.
7. NoSQL: modelos de datos y sistemas de gestión de bases de datos. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste). <http://sedici.unlp.edu.ar/handle/10915/67258>.
8. Utilización de NoSQL para resolución de problemas al trabajar con cantidades masivas de datos. Róttoli, Giovanni, López Nocera, Marcelo, Pollo Cattaneo María Florencia. 2015. XVII Workshop de Investigadores en Ciencias de la Computación (Salta, 2015). <http://sedici.unlp.edu.ar/handle/10915/45514>.
9. Aspectos de ingeniería de software, bases de datos relacionales y bases de datos no relacionales para el desarrollo de sistemas de software en escenarios híbridos. Luciano Marrero, Pablo Thomas, Ariel Pasini, Rodolfo Bertone, Eduardo Ibáñez, Verónica Aguirre, Verena Olsow, Fernando Tesone, Patricia Pesado. XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020). Universidad Nacional de la Patagonia Austral (UNPA). El Calafate, Santa Cruz (Mayo 2020). <http://sedici.unlp.edu.ar/handle/10915/104026>.
10. Análisis de performance en Bases de Datos NoSQL y Bases de Datos Relacionales. Luciano Marrero, Verena Olsow, Fernando Tesone, Pablo Thomas, Lisandro Delia y

- Patricia Pesado. XXVI Congreso Argentino de Ciencias de la Computación (CACIC 2020). Universidad Nacional de La Matanza, del 5 al 9 de octubre del 2020. <https://cacic2020.unlam.edu.ar/es-ar/>
11. Aspectos de ingeniería de software y bases de datos para el desarrollo de sistemas de software en escenarios híbridos. Patricia Mabel Pesado, Rodolfo Bertone, Pablo Thomas, Ariel Pasini, Luciano Marrero, Eduardo Ibáñez, Alejandra Rípodas, Verónica Aguirre, Verena Olsowy, Fernando Tesone. XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019). Universidad Nacional de San Juan (UNSJ). (Abril 2019). <http://sedici.unlp.edu.ar/handle/10915/77088>.
 12. Un estudio comparativo de bases de datos relacionales y bases de datos NoSQL. Pesado Patricia Mabel, Thomas Pablo, Delia Lisandro, Marrero Luciano, Olsowy Verena, Tesone Fernando, Fernández Juan Sosa. XXV Congreso Argentino de Ciencias de la Computación (CACIC 2019). Universidad Nacional de Río Cuarto, Córdoba, 14 al 18 de octubre de 2019. ISBN 978-987-688-377-1. <http://sedici.unlp.edu.ar/handle/10915/91403>.
 13. NoSQL A Brief Guide To The Emerging World of Polyglot Persistence. Pramod J. Sadalage y Martin Fowler. Pearson Education. 2013. ISBN: 978-0-321-82662-6.
 14. Data Modeling with NoSQL Database. Ajit Singh, Sultan Ahmad. ISBN 978-1072978374 (2019).
 15. Aplicaciones para Dispositivos Móviles. Estrategias y enfoques de desarrollo. Thomas Pablo Javier, et. al.. XXI Workshop de Investigadores en Ciencias de la Computación (San Juan 2019). ISBN: 978-987-3984-85-3
 16. Which Change Sets in Git Repositories Are Related?. Ramadani, J., Wagner, S. EEE 2016. International Conference on Software Quality, Reliability and Security (Viena, Austria.)
 17. Synchronization and replication in the context of mobile applications. STAGE, A. (2005 Joint Advanced Student School Course 6: Next-Generation User-Centered).
 18. A Performance Optimization Scheme for Migrating Hive Data to Neo4j Database. Publicado en 2018 International Symposium on Computer, Consumer and Control (IS3C). ISBN: 978-1-5386-7036-1. <https://ieeexplore.ieee.org/document/8644938>.
 19. III-LIDI: <http://weblidi.info.unlp.edu.ar/wp/proyectos/investigacion/>