

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^{ID}, Pablo Thomas^{ID}, Ariel Pasini^{ID}, Rodolfo Bertone^{ID}, Eduardo Ibáñez^{ID},
Verónica Aguirre^{ID}, Verena Olsowy^{ID}, Fernando Tesone^{ID}, Patricia Pesado^{ID}.

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

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

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, de nuevas aplicaciones Web, que exigen tiempos de respuestas inmediatos y generan grandes volúmenes de información. Bajo esta demanda, las bases de datos no relacionales (NoSQL), aparecen como un complemento y/o alternativa a las bases de datos relacionales en el proceso de diseño de la información. Además, se tiene como objetivo principal generar metodologías y prácticas de Ingeniería de Software considerando las características de este nuevo contexto. 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.

Palabras claves: Ingeniería de Software, Almacenamiento no estructurado, Bases de Datos NoSQL, 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.

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 [15].

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 y Bases de Datos NoSQL, con énfasis en los escenarios híbridos y las nuevas aplicaciones Web adaptables a cualquier dispositivo informático.

La disponibilidad de la información en todo momento, debido a la expansión del acceso a internet y el uso de los dispositivos móviles, han cambiado los aspectos en que la información debe ser analizada y diseñada. Se presentan así, nuevas alternativas y desafíos en la Ingeniería de Software y las arquitecturas tradicionales para el almacenamiento de la información. En este contexto, las bases de datos no relacionales representan la evolución en el almacenamiento de los datos en una nueva generación de sistemas

Web que deben responder eficientemente a las exigencias del usuario [1, 2, 8, 9 y 11]. El aumento en el acceso y procesamiento de la información digital a través de una variedad de dispositivos electrónicos con diferentes características ha alterado el concepto de diferentes ámbitos, por ejemplo, ciudades inteligentes, enseñanza virtual, domótica, etc. [10 y 11].

A razón de esto, los sistemas de software han aumentado sus exigencias, no solo el volumen de datos a procesar y almacenar, sino que también, deben responder de manera óptima al aumento de usuarios que lo acceden de forma simultánea. Esto significa que la escalabilidad y el rendimiento se han convertido en auténticos retos para la representación en el almacenamiento de la información. Las tecnologías de bases de datos más comúnmente utilizadas, las relacionales (SQL), pueden no ser suficientes y requieren de un complemento o alternativa para satisfacer las demandas actuales.

En respuesta a esta problemática, la industria del software ha incorporado a las bases de datos no relacionales (NoSQL). NoSQL no es tipo de base de datos, sino que es un conjunto de tipos de bases de datos que brinda diferentes formas para el almacenamiento no estructurado de información, es decir, operan sin un esquema predefinido. 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) [5, 6 y 7]. NoSQL propone un sistema llamado BASE (Básicamente Disponible, Estado Suave y Consistencia Eventual) [8, 9 y 11].

El desarrollo de software, año tras año, actualiza sus prácticas tradicionales de Ingeniería de Software. Los modelos y la forma en que se gestionan sus requerimientos se vuelven cada vez más ágiles. En ciertas ocasiones, la implantación de sus datos se combina en distintas categorías de almacenamiento (relacional, documental, orientado a columnas, clave-valor o grafos). Es importante el estudio de nuevas formas de almacenamiento no estructurado y comprender el funcionamiento de las herramientas que utilizan estos tipos de almacenamiento para comprender en qué tipo de problemas son aplicables. Herramientas NoSQL como MongoDB (documental), Cassandra (orientada a columnas), Redis (clave-valor) y Neo4j (grafos) son ejemplos de una nueva generación de motores de bases de datos que surgen en respuesta a diversas problemáticas existentes en las nuevas aplicaciones Web. En el almacenamiento documental, se almacena, recupera y gestiona documentos, en donde se codifican datos bajo un formato estándar (XML, JSON, BSON, etc.). En el almacenamiento orientado a columnas, los datos se organizan por columnas y no por filas, suelen ser un híbrido con el almacenamiento relacional. En el almacenamiento clave-valor, almacena un dato bajo una clave, similar a una tabla hash y en el almacenamiento de grafos, los datos son almacenados en objetos denominados nodos y son ideales para representar las relaciones o vínculos que hay entre ellos [9].

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 [4, 5, 7, 8, 13 y 14].

La calidad en las prestaciones de los dispositivos móviles ha logrado que el acceso de miles de usuarios a ciertas aplicaciones sea constante. En esta realidad, es fundamental la combinación de distintas tecnologías para lograr el mejor rendimiento posible.

El desarrollo de nuevas aplicaciones Web para escenarios híbridos y con características de movilidad, plantea nuevos desafíos en diversas áreas de la Ciencia Informática.

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, 3, 7, 8, 9 y 12].

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 Bases de Datos no relacionales (NoSQL). MongoDB, Cassandra, Redis, Neo4j, entre otras.
- Desarrollo de casos de estudio, pruebas de comparación y rendimiento de bases de datos no relacionales y relacionales [9].
- Repositorios GIT.

- Trabajo colaborativo en Proyectos de Software.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados esperados/obtenidos se pueden resumir en:

- Capacitación continua de los miembros de las líneas de investigación.
- Analizar, comparar y adaptar las nuevas metodologías y herramientas de la Ingeniería de Software para el desarrollo del software.
- Definición de técnicas, atributos y métricas para el análisis de usabilidad de sistemas en escenarios híbridos.
- Estudio y análisis de Bases de Datos no relacionales [9].
- Analizar, comparar y determinar escenarios para los distintos tipos de almacenamiento no estructurado de información.
- Comparación y análisis de resultados para diversos casos de estudio entre Bases de Datos Relacionales y Bases de Datos No Relacionales [9].
- Definición de procesos de Gestión de Incidencias utilizando repositorios GIT.
- Análisis de metodologías para la interoperabilidad de sistemas web y aplicaciones móviles [12].

El III-LIDI realiza transferencias de Aplicaciones Web y Aplicaciones Móviles en el marco de este proyecto a distintas entidades (Universidad Nacional de La Plata, Facultad de Ciencias Exactas de la UNLP, Albergue Universitario de la UNLP, entre otras).

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. Pfleegger 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. Administración de Proyectos. Guía para el Aprendizaje. Francisco Rivera Martínez, Gisel Hernández Chávez. Prentice Hall / Pearson. 2010. ISBN: 978-607-442-620-5.
6. 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).
7. 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>.
8. 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 (Abril 2019). <http://sedici.unlp.edu.ar/handle/10915/77088>. ISBN: 978-987-3984-85-3
 9. 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, Fernandez 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>.
 10. 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.
 11. Data Modeling with NoSQL Database. Ajit Singh, Sultan Ahmad. ISBN 978-1072978374 (2019).
 12. 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
 13. Which Change Sets in Git Repositories Are Related?. Ramadani, J., Wagner, S. IEEE 2016. International Conference on Software Quality, Reliability and Security (Viena, Austria.)
 14. Synchronization and replication in the context of mobile applications. STAGE, A. (2005 Joint Advanced Student School Course 6: Next-Generation User-Centered).
 15. III-LIDI: <http://weblidi.info.unlp.edu.ar/wp/proyectos/investigacion/>