

ASPECTOS DE INGENIERÍA DE SOFTWARE, BASES DE DATOS RELACIONALES, Y BASES DE DATOS NOSQL PARA EL DESARROLLO DE SISTEMAS DE SOFTWARE HÍBRIDOS.

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 diseño de Ingeniería de Software y Bases de Datos (Relacionales y No Relacionales) ante una gran variedad aplicaciones multiplataforma utilizadas por millones de usuario simultáneamente. En la actualidad, la mayoría de las aplicaciones de software son multiplataforma[2, 3, 6]. El avance en la comunicación digital, el acceso constante a la información (internet) y el aumento en la utilización de tecnología móvil, genera que estas aplicaciones sean de uso común por millones de usuarios simultáneamente. Este contexto requiere nuevas formas de pensar los diseños de Bases de Datos en complemento con las tradicionales metodologías ágiles para el desarrollo de Software. Las Bases de Datos No Relacionales surgen como una alternativa de

solución a problemas no resueltos eficientemente por los Sistemas de Gestión de Bases de Datos tradicionales (Relacionales). El término NoSQL no responde a un sólo tipo de Bases de Datos, sino que representa un conjunto de tipos de Bases de Datos, con diferentes formas y características para representar la información [1, 9, 14]. Las metodologías tradicionales de diseño y construcción de Bases de Datos (principalmente las relacionales) han sido ampliamente desarrolladas, estudiadas y aplicadas por décadas. A pesar de los avances en los procesos de la Ingeniería de Software, en donde la documentación ha perdido protagonismo, estas metodologías continúan siendo un eslabón importante en la historia de las Base de Datos, como así también, para recabar requerimientos del usuario y del contexto.

La tecnología móvil impulsó aspectos que hace algunos años no eran considerados en el desarrollo del Software, tales como, movilidad, geolocalización, generación de grandes

volúmenes de información y la diversidad de los dispositivos electrónicos involucrados [1, 2, 5, 8].

Palabras claves: Diseño de Base de Datos, Metodologías ágiles, Bases de Datos Relacionales, Bases de Datos NoSQL, Bases de Datos como servicio en la nube, Aplicaciones Multiplataforma, Aplicaciones Móviles.

CONTEXTO

Esta línea de Investigación forma parte del Proyecto “*Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital*” del Instituto de Investigación en Informática LIDI de la Facultad de Informática, 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

“*Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la*

sociedad actual: educación, ciudades inteligentes y gobernanza digital” 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: Ciudades Inteligentes sostenibles (CIS). Gobernanza Digital. Buenas prácticas y calidad*” y “*SP3: Creación de tecnologías digitales para el escenario educativo*”. [22].

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 No Relacionales), Diseño de Base de Datos (Relacionales y No Relacionales), aplicaciones multiplataforma, escenarios híbridos y aplicaciones móviles, entre otras temáticas [22].

En la actualidad existen un conjunto de nuevos desafíos para la Ingeniería de Software y las Bases de Datos. Estos desafíos van desde las etapas de diseño y perduran hasta la puesta en producción de un producto de Software.

Generalmente, un proceso de diseño para una Base de Datos, inicia con el análisis de una especificación de requerimientos en lenguaje natural, la cual deriva de la mirada de expertos sobre una problemática real. La especificación generada, posee la ambigüedad propia del lenguaje, lo que introduce un cierto grado de subjetividad al momento de plantear o pensar en un esquema de datos. Los principios o reglas que se aplican a un esquema de datos Relacional generalmente no resultan apropiados

para una Base de Datos No Relacional (NoSQL). En general, NoSQL se basa en admitir redundancia y desnormalización para obtener eficiencia en las consultas que impactarán sobre el sistema. El modelo relacional y NoSQL proponen visiones diferentes, por lo tanto, su diseño también tendrá que ser diferente [1, 7, 8, 18].

El usuario actual es cada vez más exigente y pretende que las aplicaciones respondan de manera óptima todo el tiempo y desde cualquier punto geográfico en cual posea acceso a Internet. Con el objetivo de satisfacer tales pretensiones, muchas empresas desarrollan aplicaciones que son mundialmente utilizadas poniendo el foco en la escalabilidad de sus recursos y enfrentando nuevos retos para satisfacer la demanda de sus usuarios [2, 14].

Actualmente, existe un amplio abanico de tecnologías, tanto en Bases de Datos (Relacionales y No Relacionales), como así también, en lenguajes para el desarrollo de todo tipo de aplicaciones de Software [10, 11, 12].

En el contexto de Base de Datos, existen distintos motores de Bases de Datos Relacionales y No Relacionales para el desarrollo de aplicaciones móviles, entre ellos, podemos encontrar a:

SQLite, es una librería que implementa un motor de Base de Datos Relacional autocontenido (embebido). Tiene Licencia Public Domain, y puede ser utilizada tanto en el desarrollo nativo de aplicaciones (Android o iOS), como también en aplicaciones multiplataformas [10, 14].

Interbase (Relacional), es embebido, con licencia comercial y conforme al estándar SQL. Se puede utilizar en el desarrollo nativo en Android e iOS [10].

SQLBase (Relacional), tiene licencia de uso comercial. Se encuentra disponible para el desarrollo de aplicaciones móviles nativas, en Android e iOS [10, 14, 15].

Couchbase Lite (No Relacional, Documental), es un DBMS embebido, utiliza JSON como formato de los documentos. Tiene licencia dual, y es posible utilizarlo para el desarrollo de aplicaciones móviles nativas Android e iOS, y multiplataforma [10, 12, 13, 21].

Firestore Realtime Database (No Relacional, Documental), alojado en la nube, almacena los datos en un único JSON, y cuenta con sincronización de datos en tiempo real, manteniéndose disponibles aún sin conexión. Se encuentra disponible para el desarrollo de aplicaciones nativas en Android e iOS, y en el desarrollo de aplicaciones móviles [10, 15].

Google Cloud Firestore (No Relacional, Documental), es un DBMS alojado en la nube, almacena los datos en documentos JSON, cuenta con sincronización de datos en tiempo real y mantiene los datos disponibles aún sin conexión. Es posible utilizarlo en el desarrollo de aplicaciones móviles nativas Android e iOS, y en el desarrollo de aplicaciones móviles.

Oracle Berkeley DB (No Relacional, Clave-Valor), es una familia de Bases de Datos de Clave-Valor embebidas. Tiene licencia open source y se encuentra disponible para el desarrollo de aplicaciones nativas Android e iOS [10, 17, 18].

Como parte de las investigaciones realizadas, se plantea el diseño de Bases de Datos No Relacionales. Generalmente, cuando se trabaja con Bases de Datos que poseen almacenamiento no estructurado de datos, se plantea un diseño directamente de su estructura

a nivel físico teniendo en cuenta el problema a resolver y la flexibilidad de los esquemas. Realizar un esquema a nivel conceptual para una Base de Datos NoSQL es algo que aún es tema de discusión. Otra temática de investigación se obtiene de analizar el proceso de despliegue del Software. 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 [4, 17, 18]. En este contexto, la comunicación y la sincronización entre los equipos de trabajo continúan siendo un pilar fundamental para el éxito de un proyecto. La utilización de repositorios de información y sincronización 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, 4, 6, 16].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

- Procesos de diseño para Bases de Datos No Relacionales [1, 2].
- Investigar distintos tipos de DBMSs para aplicaciones móviles [8, 10].

- Metodologías y técnicas de la Ingeniería de Software y su aplicación en el desarrollo de Software para escenarios híbridos.
- Desarrollo de casos de estudio y métricas de performance para distintos tipos de motores de Bases de Datos No Relacionales (NoSQL) como por ejemplo MongoDB, Apache Cassandra, Redis, Neo4j, entre otros[14].
- Desarrollo de casos de estudio y métricas de performance para Bases de Datos en la nube. Cloud Firestore (Google), MongoDB Atlas (MongoDB), DataStax Astra (Apache Cassandra), entre otras.
- Investigar, desarrollar casos de estudio e incorporar nuevas pruebas para la escalabilidad horizontal de motores de Bases de Datos No Relacionales.
- Investigar nuevos tipos de motores de Bases de Datos, por ejemplo: NewSQL y Bases de Datos de Serie Temporales.
- Repositorios para la sincronización del trabajo en equipo.
- Modelo de procesos para el despliegue / puesta en producción de sistemas de software

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 procesos de diseño para Bases de Datos No Relacionales.

- Estudio y análisis de distintos DBMS Bases de Datos para el desarrollo de aplicaciones móviles [9, 10].
- Estudio y análisis sobre la sincronización de requerimientos no funcionales para una Base de Datos para aplicación móvil con una Base de Datos backend [9, 10, 11].
- Estudio y análisis de Bases de Datos en la nube y otros tipos de Bases de Datos como NewSQL, Series Temporales, entre otras [10, 11, 12].
- Análisis de métricas para diversos casos de estudio realizado entre distintos tipos de motores de Bases de Datos.
- Estudio y análisis sobre escalamiento horizontal para pruebas de performance entre distintos motores de Bases de Datos [13, 15, 20].
- Definición de procesos de Gestión de Incidencias utilizando repositorios GIT [11, 12, 20, 19].
- Análisis de metodologías para la interoperabilidad de sistemas web y aplicaciones móviles [11, 12].

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 Sistemas de Software y 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, Sistema de gestión de asistencia de alumnos mediante QR, 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. Un estudio de procesos de diseño de bases de datos NoSQL. Marrero Luciano, Osowy Verena, Tesone Fernando, Thomas Pablo Javier, Corbalán Leonardo César, Fernández Sosa Juan Francisco, Pesado Patricia Mabel. XXVIII Congreso Argentino de Ciencias de la Computación (CACIC) (La Rioja, 3 al 6 de octubre de 2022). ISBN: 978-987-1364-31-2. Páginas: 404-414.
2. Aspectos de ingeniería de software, bases de datos relacionales, no relacionales y como servicios en la nube para el desarrollo de sistemas de software híbridos. Marrero Luciano, Thomas Pablo Javier, Pasini Ariel

- Cristian, Bertone Rodolfo Alfredo, Ibañez Eduardo Javier, Aguirre Verónica, Panizzi Marisa Daniela, Olsowy Verena, Tesone Fernando, Pesado Patricia Mabel. XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022, Mendoza). ISBN: 978-987-48222-3-9. Páginas: 370-376.
3. Ingeniería de Software. Un Enfoque Práctico. Novena Edición. Roger S. Pressman, Bruce R. Maxim. McGraw-Hill Interamericana Editores, S.A. 2021. ISBN: 978-1-4562-8772-6.
 4. Ingeniería de Software Teoría y Práctica. Pfleegger Shari Lawrence. Pearson / Prentice Hall. 2002. ISBN: 9789879460719
 5. Ingeniería de Software. Novena Edición. Ian Sommerville. Addison Wesley / Pearson. 2011. ISBN: 978-607-32-0603-7
 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. Model Driven Extraction of NoSQL Databases Schema: Case of MongoDB. Brahim, A.; Ferhat, R. and Zurfluh, G. (2019). In Proceedings of the 11th International Joint Conference on Knowledge Discovery. V.1: KDIR, ISBN 978-989-758-382-7, pages 145-154.
 8. *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 (UNNE). <http://sedici.unlp.edu.ar/handle/10915/67258>
 9. Análisis de performance en Bases de Datos NoSQL y Bases de Datos Relacionales. Luciano Marrero, Verena Olsowy, 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/>
 10. Un Análisis Experimental de Sistemas de Gestión de Bases de Datos para Dispositivos Móviles. 2021. Tesone Fernando, Thomas Pablo, Marrero Luciano, Olsowy Verena, Pesado Patricia. XXVII Congreso Argentino de Ciencias de la Computación (CACIC) (Modalidad virtual, 4 al 8 de octubre de 2021). Universidad Nacional de Salta (UNSA), Argentina. <http://sedici.unlp.edu.ar/handle/10915/130353>
 11. 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 Ibañez, Verónica Aguirre, Verena Olsowy, 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>.
 12. 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. Marrero Luciano, Thomas

- Pablo, Olsow Verena, Pasini Ariel, Bertone Rodolfo, Ibañez Eduardo, Aguirre Verónica, Panizzi Marisa, Pesado Patricia. XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021, Chilecito, La Rioja).
13. Aspectos de ingeniería de software y bases de datos para el desarrollo de sistemas de software en escenarios híbridos. Patricia Pesado, Rodolfo Bertone, Pablo Thomas, Ariel Pasini, Luciano Marrero, Eduardo Ibañez, Alejandra Rípodas, Verónica Aguirre, Verena Olsow, 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>.
 14. Un estudio comparativo de bases de datos relacionales y bases de datos NoSQL. Pesado Patricia, Thomas Pablo, Delía Lisandro, Marrero Luciano, Olsow 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>.
 15. A. Nori, “Mobile and embedded databases,” in Proceedings of the 2007 ACM SIGMOD International Conference on Management of data, pp. 1175-1177, 2007.
 16. Schema Extraction and Structural Outlier Detection for JSON-based NoSQL Data Stores. M. Klettke, U. Störl, S. Scherzinger, in: BTW conf., 2015, pp. 425–444.
 17. Inferring Versioned Schemas from NoSQL Databases and Its Applications. Diego Sevilla Ruiz, Severino Feliciano Morales, Jesús García Molina. https://link.springer.com/chapter/10.1007/978-3-319-25264-3_35 (Springer).
 18. 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
 19. Which Change Sets in Git Repositories Are Related?. Ramadani, J., Wagner, S. EEE 2016. International Conference on Software Quality, Reliability and Security (Viena, Austria.)
 20. Synchronization and replication in the context of mobile applications. STAGE, A. (2005 Joint Advanced Student School Course 6: Next-Generation User-Centered).
 21. S. Lee, “Creating and using databases for android applications,” International Journal of Database Theory and Application, vol. 5, no. 2, 2012..
 22. III-LIDI: <http://weblidi.info.unlp.edu.ar/wp/proyectos/investigacion/>