



# TESINA DE LICENCIATURA

**Título:** Estudio de rendimiento en MongoDB sobre arquitecturas centralizadas y distribuidas

**Autores:** Pablo José Soldi

**Director:** Prof. Fernando G. Tinetti

**Asesor profesional:** Lic. Franco Agustín Terruzzi

**Carrera:** Licenciatura en Sistemas

## Resumen

La tesina de grado tiene como objetivo principal investigar de manera experimental las facilidades de replicación y distribución de datos de una base de datos NoSQL en particular, siendo MongoDB el seleccionado.

Como caso de estudio se ha elegido el almacenamiento de datos que no son totalmente estructurados, de paso para aprovechar la experimentación con datos NoSQL. A través de la escritura y lectura de los datos, se busca obtener tiempos de respuesta y a su vez tener chequeos de consistencia. Se espera que los diferentes experimentos provean información relevante tanto para la instalación como para la operación de estas formas de almacenamiento y recuperación de datos.

La experimentación parte de una primera estructura centralizada, luego pasa a una estructura centralizada con réplicas y finaliza en una distribuida.

## Palabras Claves

Bases de datos distribuidas - NoSQL - MongoDB - Réplica - Fragmentación - Tiempos de respuesta - Consistencia

## Trabajos Realizados

Se creó un manual detallando todos los pasos para la configuración, instalación y puesta en marcha de las estructuras distribuidas que van a analizarse de MongoDB.

Se creó una aplicación web que permite unificar todas las funcionalidades necesarias para experimentar sobre MongoDB.

## Conclusiones

Las bases de datos implementadas en arquitecturas distribuidas permiten mejorar los tiempos de respuesta de una centralizada. Mientras que comparando entre estructuras distribuidas se puede mejorar en prácticamente un 50% los tiempos de respuesta si se tienen tres shards en lugar de dos.

MongoDB demostró ser consistente para escribir y leer en el manejo de grandes volúmenes de datos. Por otro lado su configuración inicial para las distintas estructuras requiere de una curva de aprendizaje grande para su uso.

## Trabajos Futuros

Resulta interesante investigar qué otras opciones además de Mongo se pueden utilizar para almacenamiento distribuido de documentos y analizar entre ellas similitudes y diferencias.