


Almacenamiento en Discord

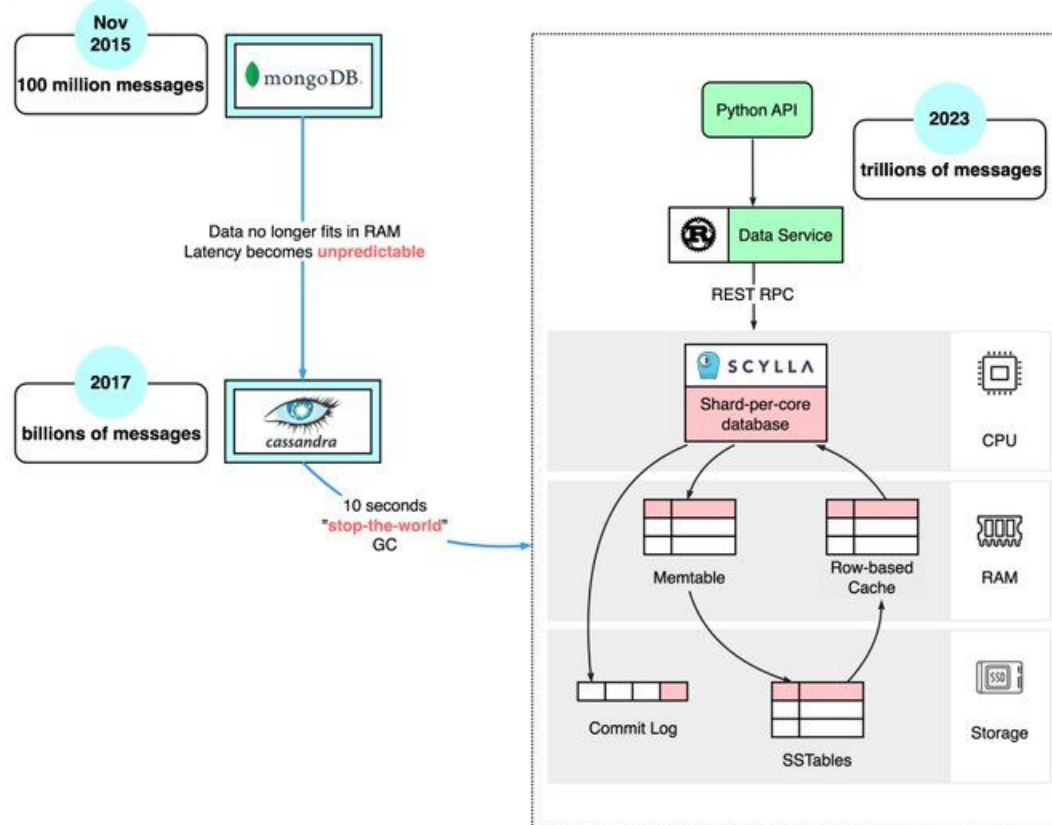
¿cómo almacena un millon de millones (trillion) de mensajes?

Evolución de la solución

- 2015: MongoDB
- 2017: Cassandra
- 2023: ScyllaDB

How Discord Stores Trillions Of Messages

 blog.bytebytego.com



2015 - MongoDB

En 2015, la primera versión de Discord se creó sobre una única réplica de MongoDB. Alrededor de noviembre de 2015, MongoDB almacenó 100 millones de mensajes y la memoria RAM ya no podía contener los datos y el índice. La latencia se volvió impredecible.

- ¿cuál era la versión disponible de MongoDB en 2015?
- ¿creeís que ahora soportaría una mayor cantidad de mensajes que 100 millones?
- ¿sabríais concretar las mejoras en Hardware desde 2015 hasta nistro días donde hacen hincapié?
- ¿pensáis que alguna vez os enfrentaréis a un proyecto con estas volumetrías?
- ¿están las organizaciones preparadas para administrar/mantener/licenciar/dar cobertura legal a estos nuevos paradigmas?

2017 - Cassandra

En **2017**, Discord tenía 12 nodos Cassandra y almacenaba miles de millones de mensajes.

A principios de **2022**, tenía 177 nodos con miles de millones de mensajes. En este punto, la latencia era impredecible y las operaciones de mantenimiento se volvieron demasiado costosas.

→ Contar batallas de la administración de Cassandra en Endesa: 24 nodos y muchos dolores de cabeza

Hay varias razones para el problema:

- 1) Cassandra utiliza el árbol LSM para la estructura de datos interna. Las lecturas son más caras que las escrituras. Puede haber muchas lecturas simultáneas en un servidor con cientos de usuarios, lo que genera puntos de acceso.
- 2) El mantenimiento de clústeres, como la compactación de SSTables, afecta el rendimiento.
- 3) Garbage Collector provoca picos en los tiempos de respuesta

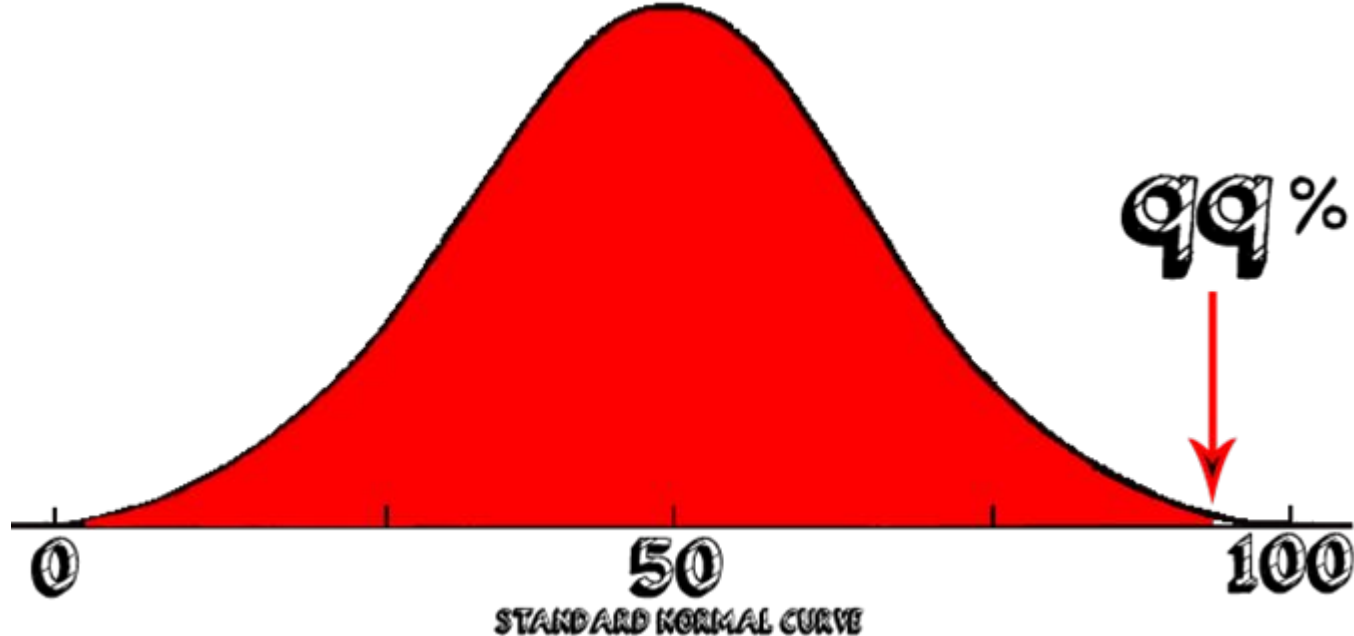
2023 - El salto a Scylla

ScyllaDB es una base de datos compatible con Cassandra, que está escrita en C++.

- Los datos se pueden migrar “fácilmente” → desarrollo adhoc de 3 meses
- Discord rediseñó su arquitectura para tener una API monolítica, un servicio de datos escrito en Rust y almacenamiento basado en ScyllaDB.
- ¿alguno podría estimar el esfuerzo en horas/hombre que tiene rediseñar una plataforma por un cambio en la capa de persistencia?
 - ¿alguien imagina el nivel de presión que tuvieron los arquitectos al tener que empujar un 3º cambio de bases de datos? ¿eso como se le vende a un jefe? ¿al que te paga la nómina todos los meses?

Algunos datos

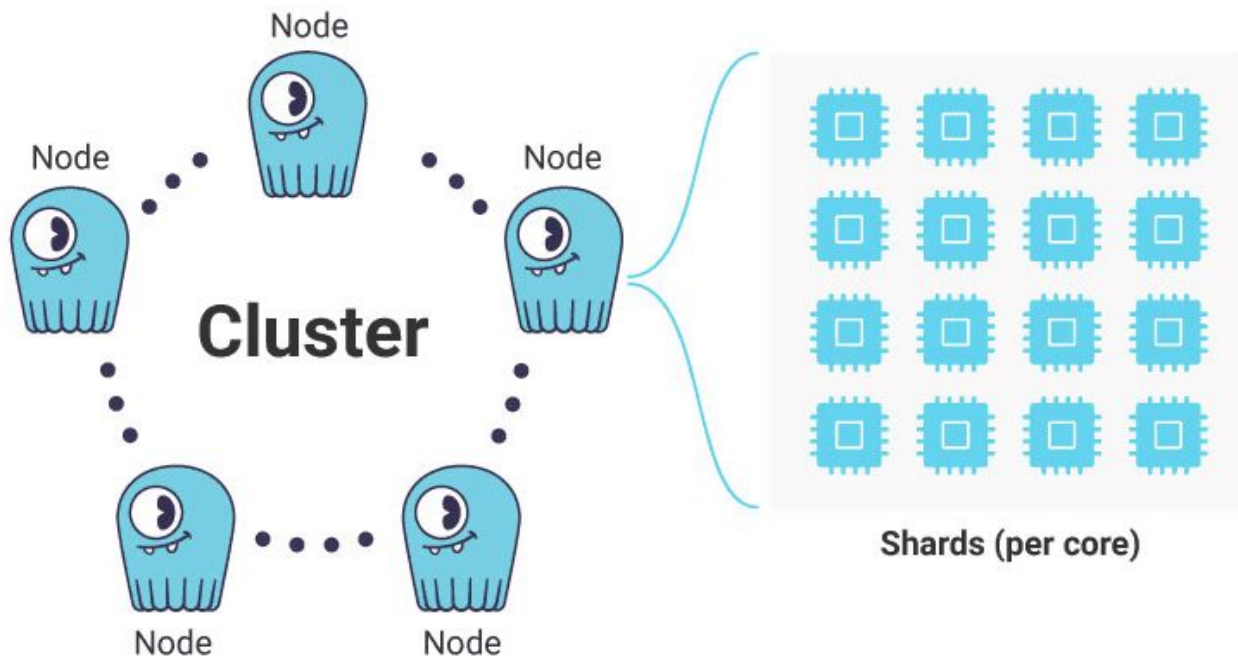
La latencia de lectura del percentil 99 en ScyllaDB es de **15 ms** en comparación con los 40-125 ms de Cassandra.
La latencia de escritura en el percentil 99 es de **5 ms** en comparación con los 5-70 ms de Cassandra.



Algunos datos

Una arquitectura optimizada para el problema

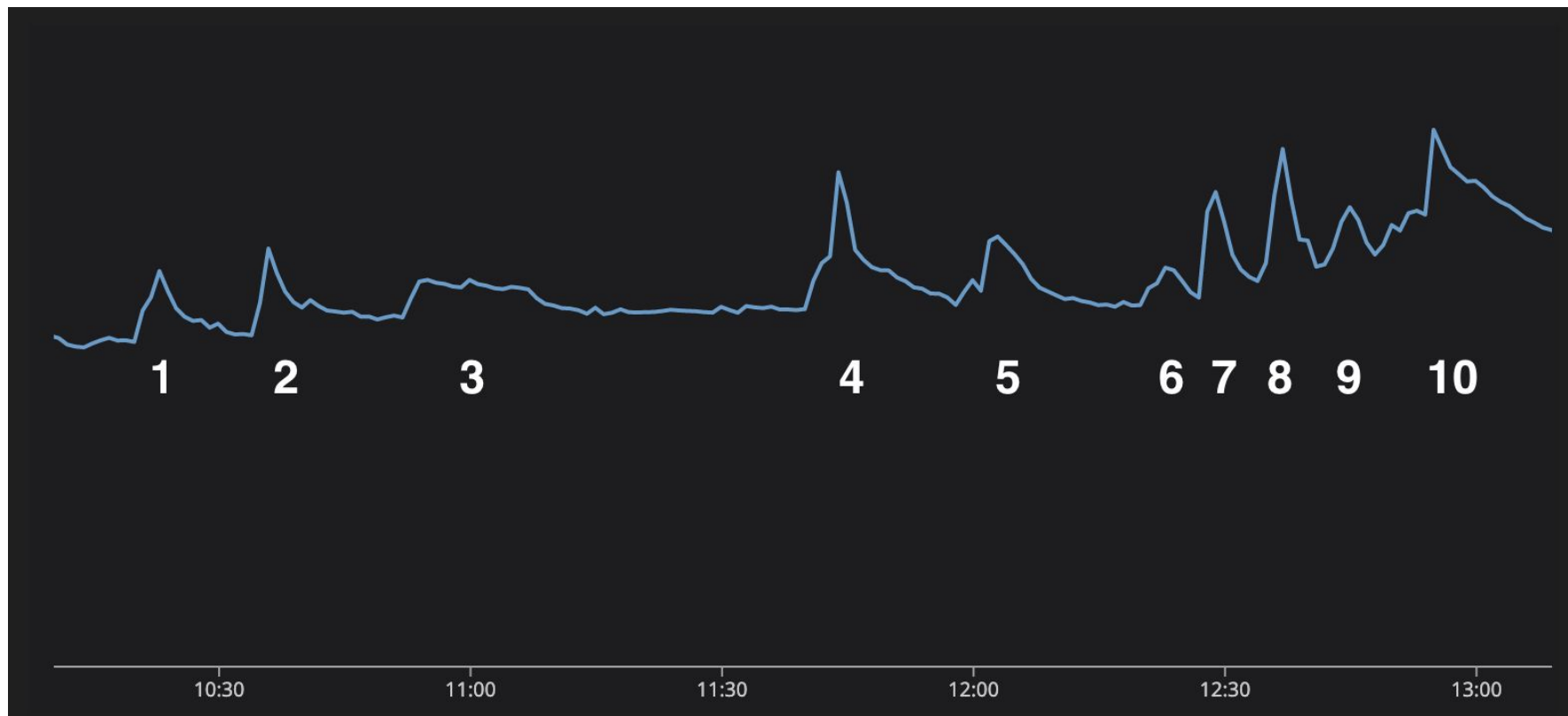
<https://www.scylladb.com/product/technology/shard-per-core-architecture/>



Más detalles - La prueba real



Más detalles - 10 momentos cumbres del partido

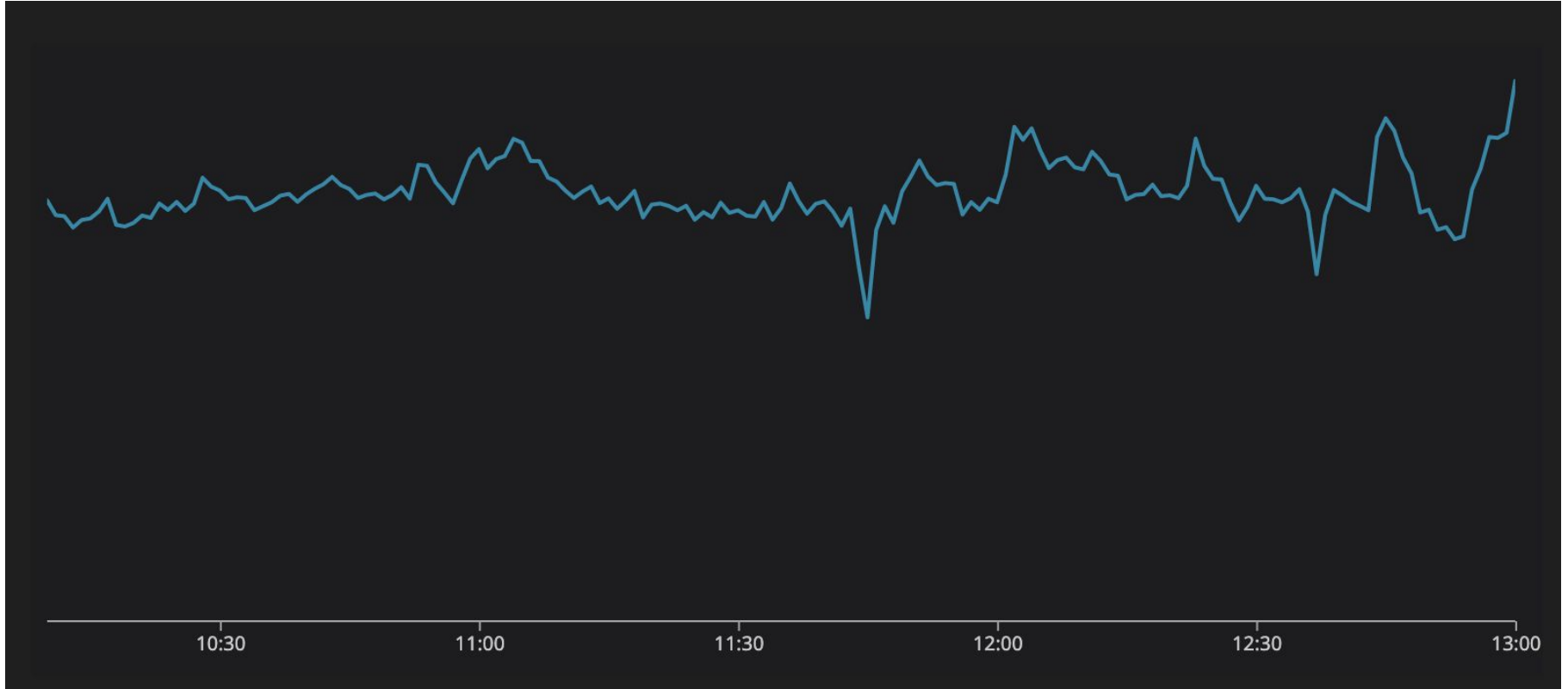


Más detalles - La prueba real

Cada uno de los diez picos en el gráfico anterior representa un evento en el partido.

1. Messi mete un penalti, y Argentina gana 1-0.
2. Argentina vuelve a marcar y gana 2-0.
3. Es el medio tiempo. Hay una meseta sostenida de quince minutos mientras los usuarios hablan sobre el partido.
4. ¡El gran pico aquí se debe a que Mbappe anota para Francia y vuelve a marcar 90 segundos después para empatar!
5. Es el final del tiempo reglamentario y este gran partido se va a la prórroga.
6. No pasa mucho en la primera mitad de la prórroga, pero llegamos al descanso y los usuarios están charlando.
7. ¡Messi vuelve a marcar y Argentina toma la delantera!
8. ¡Mbappé contraataca para empatar!
9. ¡Es el final de la prórroga, nos dirigimos a los penaltis!
10. ¡La emoción y el estrés crecen a lo largo de la tanda hasta que Francia falla y Argentina no! ¡Argentina gana!

Más detalles - Mensajes fusionados por segundo



Retro

1. Cuando te cuentan esto los propios de Discord te dicen: ¡¡Estamos contratando gente!!
 - a. ¿pensais que es una estrategia empresarial para atraer a los mejores o realmente a los más apasionados con la temática?
 - b. ¿quienes de vosotros os posicionaríais para una puesto en Discord?
2. Habrá quien se pueda creer que todos los días se toman decisiones de este calado (espero que pocos), pero yo os pregunto
 - a. ¿a quién le gustaría estar en tomas de decisiones así? ¿quien se animaría si fuera bien acompañada por un gran equipo? ¿lo haríais sin respaldo del negocio (alineacion TI - Business)
 - b. Las responsabilidades, ¿son siempre compartidas/colegiadas? ¿debe haber un voto de calidad? ¿que metodología creéis que usan para hacer estos trabajos? → os recomiendo ver la serie de Prime sobre Spotify “The Playlist”
3. Alguno pensará que esto está muy lejos de la “vida real” en Sevilla, o incluso en Madrid, Bcn o Bruselas, pero la realidad es que surgen muchas oportunidades en el día a día para alcanzar la seniority: ¡¡¡HAY QUE SABER BUSCARLAS Y TRABAJARLAS!!!

Bibliografia

- <https://www.scylladb.com/>
- <https://twitter.com/alexxybyte/status/1635306946034294784?s=20>
- <https://discord.com/blog/how-discord-stores-trillions-of-messages>
- <https://www.scylladb.com/product/technology/shard-per-core-architecture>

