

Seminar 4: Sharding Pinterest - <https://medium.com/pinterest-engineering/sharding-pinterest-how-we-scaled-our-mysql-fleet-3f341e96ca6f>

An article about how the engineers at Pinterest have scaled their MYSQL fleet.

Keywords: sharding, replication, failover, *whatever word you didn't understand* etc.

- **What is the article about?**

In general, this article discusses a data storage scaling strategy.

- **What requirements and design philosophies influenced the final solution and What is the relation between MySQL instance, database and shard?**

The system has to be scalable, robust, operate with little overhead, be highly accessible, and update with the greatest effort possible. They built many db servers running one SQL instance apiece, which would have data scattered throughout using index ranging, because a MySQL instance is essentially a single database and was a scalability bottleneck. They'd divide tables into shards and store them on separate servers. They use a mix of data type id, id in local table, and shard ID to create a UUID that closely mimics an IP address.

- **What was ZooKeeper used for?**

They utilize ZooKeeper to map a shard ID to a database, where they may search for the table with the data type id and the row with the local id.