

**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?**

This article is about 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, without operational overhead, readily accessible, and update with best effort. They built numerous 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 it was a scalability bottleneck. They'd divide tables into shards, each of which would be kept on a separate server. They've implemented UUID in such a way that it mimics an IP address - a mix of data type id, id in local table, and shard ID.

- **What was ZooKeeper used for?**

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