Member: Samiya

Week 2: Sharding Loopholes

Loophole 1: Load Balancing

In mongoDB java driver client MongoClient(addresses), it can pass in multiple mongos addresses as the entrance to the mongodb cluster, and can realize automatic failover, but how good is the load balancing? Open the source code to view:

[wKiom1fFK2KDiWHnAAAXWv5WbtE284.png](https://s3.51cto.com/wyfs02/M00/86/9B/wKiom1fFK2KDiWHnAAAXWv5WbtE284.png)

Its mechanism is to choose a machine with the fastest ping as the entry point for all requests. If this machine hangs up, the next machine will be used. In the event of Double Eleven, all requests are sent to this machine in a centralized manner, this machine is likely to hang. Once it hangs, the request will be transferred to the next machine according to its mechanism, but the total amount of pressure has not been reduced!

Loophole 2: Backup user

Once you add a user to the individual shards, which you indicate you have done for MMS, you must then have valid credentials to connect for any purpose, including mongodump. Up until you added that user for MMS, the shards were running with authentication enabled but with no users populated (this only happens if all your users are in the admin database ans using delegated auth for other databases, otherwise with 2.4 and below you would have at least one shard with users for each database - 2.6+ changes this behavior) and so you were able to connect without credentials.

Essentially this is a loophole left open so that you don't accidentally lock yourself out of your instances when you turn on auth with no users (and one that would probably have stopped working at some point as default security is tightened anyway).