

CS554 Project Ideas

ZHT:ZST – Enhance ZHT through Range Queries and Iterators

Overview

ZHT is a zero-hop distributed hash table, which has been tuned for the requirements of high-end computing systems. ZHT aims to be a building block for future distributed systems, such as parallel and distributed file systems, distributed job management systems, and parallel programming systems. The goals of ZHT are delivering high availability, good fault tolerance, high throughput, and low latencies, at extreme scales of millions of nodes. At this moment ZHT have achieved most of these goals. But it still has limitation. Because of the hash map structure which is used in ZHT, it doesn't support iteration operation and range query. Your work will be adding these features to ZHT by using sorted map (or other sorted data structure) instead of hash map and implementing corresponding operations while maintaining efficient time complexity. The modifications might not be constrained to the ZHT project, and they might have to be done in lower layers (e.g. NoVoHT).

Relevant Systems and Reading Material

ZHT paper: <http://datasys.cs.iit.edu/projects/ZHT/ZHT-CRC-PID2666213-Final.pdf>

Project URL: <http://datasys.cs.iit.edu/projects/ZHT/index.html>

Methodology

In C++, `std::map` should be the data structure that you want to use. The hard part is that you need to maintain the same iteration across different nodes.

Preferred/Required Skills

Required: Linux, C/C++ (no OOP skill needed)

Preferred: Shell scripting (for experiments)

Evaluation and Metrics

Latency, throughput, scalability

Project Mentor

Tonglin Li, tli13@hawk.iit.edu, <https://sites.google.com/site/tonglinlihome/>