Paper review [1]

Yu SU 55323707

1. The Problem

Because the performance of disk improved not very quickly. It even cannot meet the needs of large-scale web application, so the paper proposed a new class of storage called RAMCloud, where information is kept entirely in DRAM and large-scale system are created by aggregation the main memories of thousands of commodity servers. RAMCloud could provide simplification by eliminating many scalability issues, low latency and the scalable storage substrate needed for cloud computing.

2. Challenge

There are many challenges, first is reducing the latency, because it is hard to reduce the latency between network switches. And it also needs high level of durability and availability as the disk-based system and scalability for change the size of the system. And the most challenging issues is how to handle the simultaneous request interactions, because the traditional ACID model has poor scalability. Therefore, it is essential to figure out one approach which could provide great scalability for concurrency.

3. Key Insight

Overall, the paper proposed the RAMCloud, which offered a new way for organizing storage servers. It provides durable storage on the memory and achieved 100-1000 times better performance than disk-based system.

In detail, RAMCloud solved many challenging issues.

For reducing latency in network communication, on the one hand, the RAMCloud servers make good use of multi-core architectures like parallel handling. On the other hand, it use network interfaces that can be mapped directly into the application's address space by modifying the TCP or UDP protocol;

The system also could provide durability by replicating each object, while to save the space in memory, they also offered a buffered logging approach, which allows the backup data stored in the disk. And the system realized a quick recovery under these two mechanisms;

To realize the high scalability, they used a new data model which is not highly structured relational data model and they also allowed changing the data placement and movement;

And just because all these key sights, the system could contain a high consistency under the concurrency.

4. Limitation

The largest limitation of RAMCloud is it is very expensive in both execution and storage. The other limitation of RAMCloud is it is too restricted. Such as the data model with size and layout limitation and most of the optimization is oriented some kinds of application instead of the general case.

5. Future Work

Although the theory of the RAMClouds is good, but it is actually hard to implement, because there still are a lot of issues to be addressed, so the future work is optimization the architecture and as the same time try to construct it in practice.

[1] O. John, A. Parag, D. Erickson, K. Christos, (2009).The Case of RAMClouds: Scalable High-performance Storage Entirely in DRAM. *SIGOPS Operating System Review, Vol 43, No.4, December 2009, 92-105*