Week 13

Paper Critique

mClock: Handling Throughput Variability for Hypervisor IO Scheduling (OSDI '10)

This paper presents a novel algorithm, called mClock, for I/O resource allocation in a hypervisor. In a virtualized servers, qualifying QoS is essential to satisfy clients. To support QoS for I/O resource, mClock efficiently distribute I/O resources among the VMs, which are competing for the sharing resources, it schedules I/Os based on shares, reservation and limits at a per-VM level.

The advantage of suggested paper is that it guarantees isolation of the resources between VMs but they do not address the performance interference among VMs caused by sync operation which is crucial since it affects the performance directly.

Since many application calls sync operation to keep their data consistent while suffering significant performance degradation. This operation interferes other VMs directly, so efficiently isolating each VMs from affecting each other would be helpful.

FlashFQ: A Fair Queueing I/O Scheduler for Flash-based SSDs (ATC '13)

This paper suggests a novel I/O scheduler, called FashFQ, that achieves fairness and high responsiveness in the environment using flahs-based solid-state disks(SSDs). It enhances the start-time fair queueing schedulers with throttled dispatch to exploit I/O parallelism in Flash-based disk without losing fairness. Moreover, it anticipates I/O to minimize fairness violation due to deceptive idleness.

The contribution of this paper is that it suggests efficient I/O scheduler that exploits full benefit of parallelism existing in SSDs without losing the fairness between applications. However, recent research suggests that even the software stack between the application and the storage device is a overhead due to the high performance storage device. Therefore, consideration of multi-queue is necessary.

Since recent research suggests an scheduler that directly issues I/Os from application to the storage while keeping the fairness, there is an opportunity to consider the NVMe storage devices and consider it is an multi for better parallelism.