Data Management Lab
Reutlingen University

INF Informatik
Hochschule Reutlingen

DBlab

Vortrag

Data Intensive Systems on New Storage Technologies

Robert Gottstein, MSc.
Databases and Distributed Systems Group, TU-Darmstadt

Wann: am 13.12.2012 um 13:00Uhr | Raum 9-108

Abstract

As new storage technologies with radically different properties are appearing (Flash and Non-Volatile Memories), a substantial architectural redesign is required if they are to be used efficiently in a high-performance data-intensive system.

Multi-Version approaches to database systems (MVCC, SI) are gaining significant importance and become a dominating trend. They not only offer characteristics that meet the requirements of enterprise workloads, but also provide concepts that can effectively address the properties of new storage technologies. Yet version management may produce unnecessary random writes which are suboptimal for the new technologies.

A variant of SI called SI-CV collocates tuple versions, created by a transaction, in adjacent blocks and minimizes random writes at the cost of random reads. Its performance, relative to the original algorithm, in overloaded systems under heavy transactional loads in TPC-C scenarios on Flash SSD storage increases significantly. At high loads that bring the original system into overload, the transactional throughput of SI-CV increases further, while maintaining response times that are multiple factors lower. SI produces a new version of a data item once it is modified. Both the new and the old version are timestamped accordingly, which in many cases results in two independent (physical) update operations, entailing multiple random writes as well as in-place updates. These are also suboptimal for new storage technologies both in terms of performance and endurance.

We claim that the combination of multiversioning and append storage effectively addresses the characteristics of modern storage technologies.Snapshot Isolation Append Storage (SIAS) improves on SI and traditional "page granularity" append based storage managers. It manages versions as simply linked lists (chains) that are adressed by using a virtual tuple ID (VID). In SIAS the creation of a new version implicitly invalidates the old one resulting in an out-of-place write implemented as a logical append eliminating the need for invalidation timestamps. SIAS is coupled to an append-based storage manager, appending units of tuple versions. SIAS indicates up to 4x performance improvement on Flash SSD under TPC-C workload, entailed by a significant write overhead reduction (up to 38x). SIAS achieves better space utilization due to denser version packing per page and allows for better I/O parallelism and up to 4x lower disk I/O execution times. SIAS aids better endurance, due to use of out-of-place writes as appends and write overhead reduction. Compared to traditional page granularity appends, SIAS achives up to 85% higher read throughput and up to 38x write reduction.