September 2012

MADALGO seminar by Darius Sidlauskas, Aarhus University

Parallel Main-Memory Indexing for Moving-Object Query and Update Workloads

Abstract:

We are witnessing a proliferation of Internet-worked, geo-positioned mobile devices such as smartphones and personal navigation devices. Likewise, location-related services that target the users of such devices are proliferating. Consequently, server-side infrastructures are needed that are capable of supporting the location-related query and update workloads generated by very large populations of such moving objects.

This paper presents a main-memory indexing technique that aims to support such workloads. The technique, called PGrid, uses a grid structure that is capable of exploiting the parallelism offered by modern processors. Unlike earlier proposals that maintain separate structures for updates and queries, PGrid allows both long-running queries and rapid updates to operate on a single data structure and thus offers up-to-date query results. Because PGrid does not rely on creating snapshots, it avoids the stop-the-world problem that occurs when workload processing is interrupted to perform such snapshotting. Its concurrency control mechanism relies instead on hardware-assisted atomic updates as well as object-level copying, and it treats updates as non-divisible operations rather than as combinations of deletions and insertions; thus, the query semantics guarantee that no objects are missed in query results.

Empirical studies demonstrate that PGrid scales near-linearly with the number of hardware threads on four modern multi-core processors. Since both updates and queries are processed on the same current data-store state, PGrid outperforms snapshot-based techniques in terms of both query freshness and CPU cycle-wise efficiency.