SUMMARY:

Brian F. Cooper, Raghu Ramakrishnan, Utkarsh Srivastava, Adam Silberstein, Philip Bohannon, HansArno Jacobsen, Nick Puz, Daniel Weaver and Ramana Yerneni.
PNUTS: Yahoo!'s Hosted Data Serving Platform.
In Proc. of the VLDB Endowment, pages 1277-1288, 2008.

DATE: 8 February 2009

The objective of this paper is to present a parallel database implementation optimized for a large geographical distribution of its components. The system works by replicating all data at a record-level: each record will be replicated in each geographic region. The system maintains access behavior metadata for each such record in order to dynamically establish the master region for the record (the region that most write requests to the record originate from). All write operations are forwarded to the master in order to ensure a record level timeline consistency: Each write updates a version number of the record. This allows for different levels of consistency in read operation: There are three types of reads an application may perform: "read-any", "read-critical(version)", "read-latest". In order from left to right, the latency of these reads increases if the data in the local region is stale, as another region may need to get involved. In case of master failure, another replica may take over and table-level settings specify if writes should be allowed to occur at this point with the risk of violating timeline consistency.

Writes need to be propagated to all regions and this is done with a guaranteed delivery publisher/subscriber service called the Yahoo! Message Broker (YMB). This service guarantees delivery even if parts of it fail. Records are grouped together in so called 'tablets' which are the objects servers work with. The paper calls these servers "Storage units" and basically allows for any hardware storage device to be used. Tablet controllers are responsible for assigning tablets to storage units and performing load control by moving tablets around.

The service is live and being used in several Yahoo! Products.

SUMMARIZED BY: Cătălin-Alexandru Avram