# Efficient, Consistent and Secure Global-Scale Data Management

Sujaya Maiyya* Faisal Nawab† Cetin Sahin** Victor Zakhary*
Divyakant Agrawal* Amr El Abbadi*
*UC Santa Barbara, **SAP, †UC Santa Cruz
*{sujaya_maiyya, victorzakhary, agrawal, elabbadi}@ucsb.edu, **{cetin08@gmail.com} †{fnawab@ucsc.edu}

***Abstract*—Processing and analyzing data is becoming increasingly ubiquitous and is the driving force behind the sustained growth of Internet applications and the emergence of Big Data Analytics. These applications typically adopt the cloud model where they are hosted in a single datacenter. This introduces a fundamental limitation: communication to a centralized datacenter incurs significant latencies. The utilization of edge nodes is inevitable for the future success and growth of many emerging low latency and mobile applications. In this talk, we will explore various technologies that aim to facilitate building global-scale and edge-aware data management systems. These approaches are based on Geo-replication, where data is replicated across geographic locations to be closer to users, and Edge-awareness, where applications are deployed on edge locations to bypass the last-mile infrastructure. We propose novel consensus approaches that manage access to partitioned data across globally-distributed datacenters and edge nodes. The main objective is to reduce the latency of serving user requests, while ensuring fault-tolerance and adapting gracefully to mobility. In addition to failures, data centers are constantly exposed to an increasing number of non-trivial adversarial threats. Traditional cryptographic methods either limit the functionality of the data, or significantly increase retrieval costs. We will highlight some novel approaches that ensure efficient privacy preserving access to data in the Cloud.**

***Index Terms*—Data Management, Cloud, Edge, Privacy.**

The utilization of edge nodes is inevitable for the success and growth of emerging low latency applications, such as Augmented and Virtual Reality (AR/VR) and vehicular networks. Such applications have stringent latency requirements that the current cloud model cannot satisfy. This is due to the large communication latency between users and their closest datacenter (up to 100ms). This latency problem is exacerbated for applications that serve users across large geographical areas. In such cases, users incur wide area latency as large as 100s of milliseconds to seconds. Placing data closer to users at edge nodes overcomes this fundamental communication latency limit. We envision, as others have, that the cloud model will extend to edge locations similar to how content delivery networks utilize edge locations. However, rather than edge locations being used for data caching only, they will also host data management components that will allow manipulation and querying of local edge partitions. In this presentation, we focus on transaction processing as the data management task to be supported by the edge data management components. The model and focus of this aims to serve web and cloud applications, such as online shops, social networks, and collaborative applications.

We present Dynamic Paxos (DPaxos), a Paxos-based consensus protocol [1] to manage access to partitioned data across globally-distributed datacenters and edge nodes. DPaxos is intended to implement a State Machine Replication component in data management systems for the edge. DPaxos targets the unique opportunities of utilizing edge computing resources to support emerging applications with stringent mobility and real-time requirements such as Augmented and Virtual Reality and vehicular applications. The main objective of DPaxos is to reduce the latency of serving user requests, recovering from failures, and reacting to mobility. DPaxos achieves these objectives by a few proposed changes to the traditional Paxos protocol. Most notably, DPaxos proposes a dynamic allocation of quorums (i.e., groups of nodes) that are needed for Paxos Leader Election. Leader Election quorums in DPaxos are smaller than traditional Paxos and expand only in the presence of conflicts.

DPaxos proposes Zone-centric Quorums as an alternative to majority-based techniques to avoid unnecessary wide-area communication. A zone denotes a collection of neighboring edge nodes. DPaxos restricts the communication corresponding to a data partition to be within the zones where its users are located. To do this, DPaxos distinguishes between the quorums that are needed to perform the main two tasks in Paxos: Leader Election (coordination with other nodes to select a leader for a partition and is typically invoked in reaction to failures or mobility) and Replication (committing data from a leader to secondary nodes and is typically invoked for every transaction or request). In typical workloads, Replication is more frequent than Leader Election, and thus data management systems should be prioritized to optimize its performance. Ideally, for performance, Replication would be performed within a zone rather than a majority of nodes. Flexible Paxos, proposed by Howard et al. [2], shows that it is possible to assign arbitrarily small Replication quorums as long as they satisfy the condition: a Leader Election quorum must intersect all Replication quorums. This means that in Flexible Paxos-based approaches, the trade-off of small Replication quorums within zones is an expensive Leader Election quorum that must span all zones.

We base DPaxos Zone-Centric Quorums on the theoretical foundation laid by Flexible Paxos and adapt its quorum alloca-

tion techniques to the practical application of data management on globally-distributed edge nodes. Then, we propose two approaches to overcome Flexible Paxos significant Leader Election penalty:

1) Expanding Quorums: this approach overcomes Flexible Paxos intersection condition and allows both Leader Election and Replication quorums to be small. DPaxos is the first Paxos protocol that allows Leader Election to not intersect with all Replication quorums. Rather, the Leader Election quorum starts small and then grows to only intersect with Replication quorums that are being used by other leaders.
2) Leader Handoff: this approach supports fast leader mobility. Mobility, unlike failures, is triggered by known user actions, and hence can be exploited to optimize Leader Election. DPaxos exploits this to enable Leader Election via a lightweight, single round of messaging between the old and the new leaders.

In addition to failures, data centers are constantly exposed to an increasing number of non-trivial adversarial threats. Traditional cryptographic methods either limit the functionality of the data, or significantly increase retrieval costs. During the last decade, a large body of academic work has tackled the problem of outsourcing databases to an untrusted cloud while maintaining both confidentiality and SQL-like querying functionality (at least partially). We will highlight some novel approaches that ensure efficient privacy preserving access to data in the Cloud. In particular, we briefly discuss TaoStore [3] and PinedRQ‘ [4].

TaoStore is an oblivious storage systems that hides both the contents of the data as well as access patterns from an untrusted cloud provider. The target scenario is one where multiple users from a trusted group (e.g., corporate employees) asynchronously access and edit potentially overlapping data sets through a trusted proxy mediating client-cloud communication. TaoStore is built on top of a new tree-based ORAM scheme that processes client requests concurrently and asynchronously in a non-blocking fashion. This results in a substantial gain in throughput, simplicity, and flexibility over previous systems.

PinedRQ is a differentially private index for outsourced encrypted dataset and that supports non-aggregate range queries on cloud stored data, while achieving both privacy and efficiency. Performing range queries efficiently in an untrusted cloud setting has not been addressed in a satisfactory manner. Range queries express a bounded restriction over the retrieved records. They are fundamental database operations. PinedRQ sends two complementary data structures to the cloud: an encrypted version of the database, e.g., AES encryption scheme, indexed by a hierarchy of histograms, such that both are perturbed to satisfy differential privacy. Efficiency comes from the disclosure of the index, in the clear, to the cloud, for guiding the query execution strategy. No computation is ever performed on encrypted data. Privacy comes from the differential privacy guarantees of the function that computes the encrypted database and the index. Indeed, the differential privacy model is todays de facto standard for protecting personal information that needs to be partially disclosed. PINED-RQ executes queries in the order of at least one magnitude faster.

## I. Acknowledgment

This work was partially funded by the NSF grants CNS-1528178, CNS-1703560 and CNS 1815733.

## References

[1] Faisal Nawab, Divyakant Agrawal, and Amr El Abbadi. Dpaxos: Managing data closer to users for low-latency and mobile applications. In *Proceedings of the 2018 International Conference on Management of Data, SIGMOD Conference 2018, Houston, TX, USA, June 10-15, 2018*, pages 1221–1236, 2018.

[2] Heidi Howard, Dahlia Malkhi, and Alexander Spiegelman. Flexible paxos: Quorum intersection revisited. In *20th International Conference on Principles of Distributed Systems, OPODIS 2016, December 13-16, 2016, Madrid, Spain*, pages 25:1–25:14, 2016.

[3] Cetin Sahin, Victor Zakhary, Amr El Abbadi, Huijia Lin, and Stefano Tessaro. Taostore: Overcoming asynchronicity in oblivious data storage. In *IEEE Symposium on Security and Privacy, SP 2016, San Jose, CA, USA, May 22-26, 2016*, pages 198–217, 2016.

[4] Cetin Sahin, Tristan Allard, Reza Akbarinia, Amr El Abbadi, and Esther Pacitti. A differentially private index for range query processing in clouds. In *34th IEEE International Conference on Data Engineering, ICDE 2018, Paris, France, April 16-19, 2018*, pages 857–868, 2018.