Letter from the Special Issue Editor

For decades the database community has looked at *consistent data replication and distribution* – most often in the context of the transactions offering the isolation and atomicity properties. The proliferation of distributed and replicated database systems, be it within a cloud, across data centres, or in mobile environments, has led to a rich research literature offering manyfold solutions to distributed transaction management and replica control. In the March Issue 2015 of the IEEE Bulletin, I invited experts in this area, mainly from the database community, to present their views and work on data consistency in the cloud, with a focus on transaction management. In contrast, this issue aims in giving an insight into how other research communities address the challenges of data consistency as this is such a cross-cutting concern that requires solutions from many areas of computer science: distributed systems, distributed algorithms, storage systems, operating systems, computer architecture and programming language, to name a few.

While each of the articles looks at the problem with the perspective of a specific research discipline it becomes apparent that there is a common thread of main building blocks that are fundamental to providing consistency in distributed and replicated environments. Examples are well-known protocols such as consensus, state-machine replication, or 2-phase commit, correctness criteria such as serializability, linearizability, causal and eventual consistency, and properties such as scalability, availability, and throughput and response time performance. This shows the importance of a cross-disciplinary approach to data consistency.

The first set of papers sheds light on the challenge to actually understand what "consistency for replicated data" means, as there are many possible interpretations and trade-offs. In the first article, Aguilera and Terry look at the notions of consistency across disciplines. The authors identify two broad types of consistency, namely state- and operation consistency, and show how they map to the many existing definitions of consistency developed for distributed systems, database systems, and computer architecture. In the second article, Guerraoui *et al.* provide a comprehensive overview of the trade-offs between many different properties in replicated systems such as the level of data consistency offered, performance in terms of latency and/or scalability, and other aspects such as availability and robustness to failures and churn. In the next article, Zhang *et al.* offer a detanglement of the often intertwined tasks of distributed concurrency control, distributed commit, and replication protocols that offer availability and consistent data despite failures. In particular, the authors provide a detailed discussion of when and when not operations need to be ordered across all replicas. In the following article, Vukolić takes a different viewpoint, focusing on strongly consistent protocols but looks how they can be implemented not only in software, but also in hardware, and for blockchains.

From there, the next two articles look at consistency issues from a programming language perspective. Burckhardt and Protzenko describe the concept of abstract executions to specify the behaviour of a replicated system. Such a specification approach helps in investigating fundamental properties of replication solutions. Sivaramakrishnan *et al.* present Quelea, a declarative programming model for eventually consistent data stores that allows for expressing precise high-level consistency guarantees without the need to know implementation-specific low-level data store semantics.

The last three articles present propose concrete frameworks and solutions for data distribution and replication. Altınbüken and van Renesse present a software-defined framework for building large-scale distributed systems that handles dynamic change and growth by allowing components to be transformed dynamically, be it to support replication, batching, sharding or encryption. Balegas *et al.* present a solution to geo-replication that merges weak and strong consistency solutions in order to find the right balance between consistency and performance. Finally, Bezerra *et al.* looks at performance aspects of strong consistency solutions based on the state machine approach at scale.

I hope you enjoy this collection of articles as much as I did!

Bettina Kemme
McGill University
Montreal, Canada