Dial M for Management: Next Generation NoSQL

K. M. Dahlgren
University of California, Santa Cruz
kmdahlgr@ucsc.edu

1. INTRODUCTION

NoSQL databases offer a powerful and flexible means of querying non-relational data. However, leading NoSQL systems typically achieve high performance goals while minimizing support for traditional data management services and defying the establishment of solid formal models. In particular, the systems generally shun tools and design principles rendered conventional by the long history of RDBMSs, including join operations, aggregate functions, and integrity constraints. Consequently, users of NoSQL technologies are forced to find alternative means of supporting the essential missing features by either loading additional software packages over the NoSQL infrastructure or implementing the functionality themselves. While the resulting systems represent custom solutions highly optimized for specific applications, such a lack of built-in support for generally needed tools and services defies the abstraction philosophy central to principled software design [1]. Likewise, the paradigm generally flouts the need for principled design models guiding the assurance of system integrity at the theoretical level [2]. This paper articulates the problem surrounding the lax management and model provisions pursued by current NoSQL databases and outlines a number of challenges for future research to mold the next generation of NoSQL systems into more principled alternatives to traditional DBMSs.

Problem. The NoSQL paradigm predominantly ignores the potential for NoSQL databases to be highly efficient while providing built-in support for sophisticated management services and model specifications. The lack of sophisticated management services violates software quality standards by forcing users to provide standard functionality themselves. Furthermore, absent model specifications disregard the importance of conceptual integrity as a key element in assuring the validity of the data management protocols.

2. CHALLENGES

Infusing principled management in NoSQL systems presents a number of challenges.

- **Performance is power.** Traditional DBMSs provide complete packages with many built-in features, strong usability support, and reliable formal models. However, the overriding concern for the NoSQL paradigm is performance. A major challenge for next generation NoSQL is overcoming the temptation to chase the latest optimization carrot and consider sophisticated management and models as important design factors.

- **Features as second-class citizens.** NoSQL systems typically focus on finding and latching onto a novel method for optimizing a particular subset of standard access operations under certain conditions (deployment on distributed systems, in-memory processing, streaming data, etc.). Features supporting more powerful management services trickle into the product as an after-thought. Next generation systems must take the built-in implementation of essential features and services vital to many applications more seriously.

- **The "Whatever Works" model.** Unlike traditional RDBMSs, NoSQL databases avidly buck the idea of building systems around formal models grounded in solid theory [3]. Instead, performance is paramount and principles guiding the overall conceptual integrity of the system are neglected. A key challenge for next generation NoSQL is developing sound methodologies for proving and ensuring conceptual integrity.

3. CONCLUSION

The relatively short history of NoSQL technologies is marked by a massive boom in fast and flexible storage infrastructures with little to no support for traditional data processing tools, basic usability, maintainability, and safety provisions, or theoretical soundness. The circumstance violates a number of core principles related to software design quality and to the theoretical and motivational roots of the field of databases. Accordingly, several exciting challenges line the path of future research toward casting the NoSQL paradigm as a sound alternative to traditional DBMSs.

4. REFERENCES

[1] E. R. Baker. Basic principles and concepts for achieving quality, 2007.
[2] C. J. Date and H. Darwen. *Databases, Types and the Relational Model: The Third Manifesto.* Addison-Wesley Longman, 2007.
[3] V. Varga. Conceptual design of document nosql database with formal concept analysis.