

Riak DT Map: A Composable, Convergent Replicated Dictionary

Russell Brown

Basho Technologies, Inc.
russelldb@basho.com

Sean Cribbs

Basho Technologies, Inc.
sean@basho.com

Sam Elliott

Basho Technologies, Inc.
sam.elliott@basho.com

Christopher Meiklejohn

Basho Technologies, Inc.
cmeiklejohn@basho.com

Abstract

Conflict-Free Replicated Data-Types (CRDTs) [6] provide greater safety properties to eventually-consistent distributed systems without requiring synchronization. CRDTs ensure that concurrent, uncoordinated updates have deterministic outcomes via the properties of bounded join-semilattices.

We discuss the design of a new convergent (state-based) replicated data-type, the Map, as implemented by the Riak DT library [4] and the Riak data store [3]. Like traditional dictionary data structures, the Map associates keys with values, and provides operations to add, remove, and mutate entries. Unlike traditional dictionaries, all values in the Map data structure are also state-based CRDTs and updates to embedded values preserve their convergence semantics via lattice inflations [1] that propagate upward to the top-level. Updates to the Map and its embedded values can also be applied atomically in batches. Metadata required for ensuring convergence is minimized in a manner similar to the optimized OR-set [5].

This design allows greater flexibility to application developers working with semi-structured data, while removing the need for the developer to design custom conflict-resolution routines for each class of application data. We also discuss the experimental validation of the data-type using stateful property-based tests with QuickCheck [2].

Categories and Subject Descriptors C.2.4 [Distributed Systems]: Distributed databases; D.3.3 [Programming Techniques]: Language Constructs and Features - abstract data types, patterns, control structures; E.1 [Data Structures]: Distributed data structures; H.2.4 [Database Management Systems]: Distributed databases

Keywords Dynamo, Eventual Consistency, Data Replication, Commutative Operations, Riak, Erlang, Property-based Testing, QuickCheck

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

PaPEC '14, April 13-16, 2014, Amsterdam, Netherlands.
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-2716-9/14/04...\$15.00.
<http://dx.doi.org/10.1145/2596631.2596633>

Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 609551.

References

- [1] P. S. Almeida, C. Baquero, and A. Cunha. Composing Lattices and CRDTs. In B. Kemme, G. Ramalingam, A. Schiper, M. Shapiro, and K. Vaswani, editors, *Consistency in Distributed Systems (Dagstuhl Seminar 13081)*, volume 3, pages 92–126, Dagstuhl, Germany, Feb. 2013. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. URL <http://www.dagstuhl.de/mat/Files/13/13081/13081.BaqueroCarlos.Slides.pdf>.
- [2] T. Arts, L. M. Castro, and J. Hughes. Testing Erlang Data Types with Quviq Quickcheck. In *Proceedings of the 7th ACM SIGPLAN Workshop on ERLANG, ERLANG '08*, pages 1–8, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-065-4. URL <http://doi.acm.org/10.1145/1411273.1411275>.
- [3] Basho Technologies, Inc. Riak source code repository. <https://github.com/basho/riak>, 2009-2014.
- [4] Basho Technologies, Inc. Riak DT source code repository. https://github.com/basho/riak_dt, 2012-2014.
- [5] A. Bieniusa, M. Zawirski, N. Pregoça, M. Shapiro, C. Baquero, V. Balesgas, and S. Duarte. An optimized conflict-free replicated set. *ArXiv e-prints*, Oct. 2012.
- [6] M. Shapiro, N. Pregoça, C. Baquero, and M. Zawirski. A comprehensive study of Convergent and Commutative Replicated Data Types. Rapport de recherche RR-7506, INRIA, Jan. 2011. URL <http://hal.inria.fr/inria-00555588>.