



CoherentPaaS

Coherent and Rich PaaS with a
Common Programming Model

ICT FP7-611068

NoSQL data store final implementation

D5.3

September 2015

Document Information

Scheduled delivery: 30.09.2015
 Actual delivery: 30.09.2015
 Version: v1.0
 Responsible Partner: FORTH

Dissemination Level

PU Public
 PP Restricted to other programme participants (including the Commission)
 RE Restricted to a group specified by the consortium (including the Commission)
 CO Confidential, only for members of the consortium (including the Commission)

Revision History

Date	Editor	Status	Version	Changes
08-aug-2015	A. Bilas	Draft	V0.1	Table of contents
03-oct-2014	G. Saloustros	Draft	V0.2	Key-value store content
19-aug-2015	P. Kranas	Draft	V0.3	Document store (MongoDB) content
01-09-2015	R. Pau	Draft	V0.4	Graph database (Sparksee) content
08-09-2015	G. Saloustros	Draft	V0.5	Draft for internal review
18-09-2015	M. Matos	Draft	V0.6	Draft after internal review
22-09-2015	Marta Patiño	Draft	V0.7	Draft after internal review
24-09-2015	Iván Brondino	Draft	V0.8	Draft after internal review
25-09-2015	Marta Patiño	Draft	V0.9	Draft after internal review
28-09-2015	R. Pau	Draft	V0.9	Edits
28-09-2015	P. Kranas	Draft	V0.9	Edits
29-09-2015	G.Saloustros	Draft	V1.0	Final version

Contributors

Angelos Bilas (FORTH), Giorgos Saloustros (FORTH), Pavlos Kranas (ICCS), Sotiris Stamokostas (ICCS), David Dominquez-Sal (Sparsity), Raquel Pau (Sparsity)

Internal Reviewers

Ricardo Jimenez (UPM-LeanXcale), Marta Patiño (UPM) Rui Oliveira (UMinho), Miguel Matos (UMinho), (UPM), Iván Brondino (UPM)

Acknowledgements

Research partially funded by EC 7th Framework Programme FP7/2007-2013 under grant agreement n° 611068.

More information

Additional information and public deliverables of CoherentPaaS can be found at: <http://coherentpaas.eu>

1 Executive Summary

CoherentPaaS aims at integrating different SQL and NoSQL data stores in a single cloud PaaS, under a Common Query language and a scalable transactional management system across stores. This report discusses the required enhancements to the NoSQL data stores to enable their integration with the Common Query Language, CloudMdsQL, and the transactional stack.

The NoSQL data stores that are integrated into CoherentPaaS are the document data store MongoDB, the key-value data store H-Eutropia, and the graph database Sparksee. The appropriate enhancements performed for each store are:

- In MongoDB, its multiversion concurrency control mechanism is altered by storing the private write-set of each transaction into the data store storage, instead of keeping it in-memory. This created the benefit of relying in MongoDB native rich query engine for executing selections, while also reducing the memory space needed for the execution of an application. The integration with the Transactional Manager is completed and MongoDB can now handle concurrent transactions and ensure ACID semantics. Finally, MongoDB is integrated with the Common Query Engine (CQE) by implementing a wrapper that receives queries written in the CloudMdsQL language and transforms them to MongoDB's native language for execution.
- In H-Eutropia, the cloud data store resulting from the integration of Eutropia key-value data store with HBase. H-Eutropia replaces HBase storage back end with Eutropia key value store described in D5.1 targeting clusters equipped with fast storage devices. In the prototype that corresponds to this report, Eutropia is fully integrated with HBase supporting its put/get/scan API. It supports versioning, which is necessary for the integration with the transactional stack, it implements the appropriate functions for scalability, and it provides the appropriate wrappers for the integration with CQE.
- Finally, Sparksee graph database has applied the appropriate modifications to support versioning. It is now fully integrated with CoherentPaaS transactional stack and with CQE through the development of the corresponding wrapper.