
CoherentPaaS

Coherent and Rich PaaS with a Common Programming Model

ICT FP7-611068

# X-Ray Subsystem Implementation (initial version)

D8.4

September, 2015

## Document Information

| | |
|---|---|
| Scheduled delivery | 30.09.2015 |
| Actual delivery | 30.09.2015 |
| Version | 0.2 |
| Responsible Partner | INESC |

## Dissemination Level:

RE Restricted to a group specified by the consortium (including the Commission)

## Revision History

| Date | Editor | Status | Version | Changes |
|---|---|---|---|---|
| 27.08.2015 | J. Pereira | Draft | 0.1 | Initial revision. |
| 15.09.2015 | J. Pereira | Draft | 0.2 | Corrections after internal revision. |
| | | | | |
| | | | | |
| | | | | |

## Contributors

P. Guimarães (INESC TEC), J. Pereira (INESC TEC)

## Internal Reviewers

Sotiris Stamokostas (ICCS/NTUA), Ying Zhang (MonetDB Solutions)


## 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

The integration of multiple database technologies, including both SQL and NoSQL, allows using the best data management tools tools for each aspect of a complex data centered application and is increasingly sought in practice. Unfortunately, this makes it difficult for database developers and administrators to obtain a clear view of the resulting composite data processing paths, as they combine operations chosen by different query optimisers, implemented by different software packages, and partitioned across distributed systems.

In CoherentPaaS this challenge is addressed with the X-Ray Subsystem, a framework for monitoring and analysis of distributed and heterogeneous data processing systems. First, it provides a way to add monitoring code to applications and data stores running in the Java platform. By using bytecode instrumentation, this does not rely on the availability of the source code and can be applied conditionally to avoid overhead in production systems. In addition, it provides mechanisms for integration of C/C++ applications. Second, it provides mechanisms for tracking the interaction of multiple threads, on synchronisation primitives, and of distributed processes communicating with sockets.

Preliminary implementations of the analysis and visualization components are presented for completeness. These implementations already provide a proof-of-concept of the main contribution: the ability to track requests across distribution and heterogeneity boundaries in the CoherentPaaS system. They are preliminary as they limit the amount of data that can be processed, hence the scale of what can be monitored, and provide limited visualization options.

The customization of the generic subsystem for the CoherentPaaS data-stores is described in a different document [29].

The software described here is version 0.5.0, which is available in the CoherentPaaS Artifactory (both the end-user binary distribution and Maven artifacts) and git servers.