"Understanding Schema Evolution in Schema-less NoSQL Data Stores"

Loup Meurice and Anthony Cleve

PReCISE Research Center
University of Namur

Abstract:

NoSQL data stores are becoming popular due to their schema-less nature. It offers great flexibility, since it does not require to declare a global schema. Thus, the data model is maintained within the application source code. However, due to this flexibility, developers have to struggle with a growing data structure entropy and have to manage legacy data. However, support to schema evolution cruelly lacks, what may come to runtime errors or irretrievable data loss, if not properly handled. In this paper, we present a tool approach which allows understanding schema evolution of a schema-less NoSQL data store by analyzing the application source code and its history. We motivate this approach on a subject system and explain how such an approach is essential to understand the present (e.g., detecting runtime errors or data loss) and facilitate future developments.