A distributed archival network for process-oriented autonomic long-term digital preservation

Author(s) Subotic, Ivan; Rosenthaler, Lukas; Schuldt, Heiko
Author(s) at UniBasel Schuldt, Heiko ; Subotic, Ivan ; Rosenthaler, Lukas ;
Year 2013
Title A distributed archival network for process-oriented autonomic long-term digital preservation
Book title (Conference Proceedings) Proceedings of the 13th ACM/IEEE joint conference on digital libraries (JCDL '13)
Place of Conference Indianapolis, IN, USA
Publisher Association for Computing Machinery
Place of Publication New York, N.Y.
Pages S. 29-38
The reliable and consistent long-term preservation of digital content and metadata is becoming increasingly important – even though the storage media used are potentially subject to failures, or the data formats may become obsolete over time. A common approach is to replicate data across several sites to increase their availability. Nevertheless, network, software, or hardware failures as well as the evolution of data formats have to be coped with in a timely and, ideally, an autonomous way, without intervention of an administrator. In this paper we present DISTARNET, a distributed, autonomous long-term digital preservation system. Essentially, DISTARNET exploits dedicated processes to ensure the integrity and consistency of data with a given replication degree. At the data level, DISTARNETă supports complex data objects, the management of collections, annotations, and arbitrary links between digital objects. At process level, dynamic replication management, consistency checking, and automated recovery of the archived digital objects is provided utilizing autonomic behavior governed by preservation policies without any centralized component. We present the concepts and implementation of the distributed DISTARNET preservation approach. Most importantly, we provide details of the qualitative and quantitative evaluation of the DISTARNET system. The former addresses the effectiveness of the internal preservation processes while the latter evaluates DISTARNET's performance regarding the overall archiving storage capacity and scalability.

Digital Object Identifier DOI 10.1145/2467696.2467710