

Distributed digital preservation in the cloud

David S.H. Rosenthal and Daniel L. Vargas

The International Journal of Digital Curation, ISSN 1746-8256, Vol. 8, n. 1, 2013, p. 107-119

El programa LOCKSS (Lots of Copies Keep Stuff Safe) de la Biblioteca de la Stanford University fue pionero hace unos catorce años en el concepto de Preservación Digital Distribuida, mediante la construcción de una red de cajas LOCKSS en las que las bibliotecas podían recolectar y preservar el contenido publicado en la Web. Actualmente, cerca de 150 bibliotecas organizan cajas LOCKSS, preservando revistas electrónicas y libros electrónicos en la Red Global LOCKSS (GLN, en sus siglas en inglés), y bases de datos, documentos gubernamentales, colecciones especiales y otros contenidos en Redes Privadas (LOCKSS PLNs). Estas cajas, generalmente son modestos servidores de PC con grandes cantidades de almacenamiento en el disco local. Algunas bibliotecas de la red se plantearon la posibilidad de utilizar el almacenamiento asequible en la nube para sus cajas LOCKSS. Este artículo describe el experimento en el que se organizó una producción de cajas LOCKK para preservar el contenido GLN en el Servicio de nube de Amazon. En primer lugar, los autores describen las características de los servicios típicos en la nube, y luego discuten las diversas arquitecturas técnicas posibles por las cuales una caja LOCKSS podría utilizar este tipo de servicios, descartando algunas por razones de rendimiento. En el experimento implementaron una Amazon Machine Instance (AMI), respaldada por S3 y configurada mediante un “bootbucket” que especificaba los parámetros del sistema. Se detectaron problemas de rendimiento en el uso de S3 y la necesidad de una mejora de la API para hacerles frente. Se comprobó que la integridad de los contenidos en los

actuales sistemas de almacenamiento en la nube también era problemática. Esto, unido al precio del servicio de almacenamiento en la nube, mucho mayor que el del disco local, hizo que se descartara esta arquitectura. El experimento sugiere que con algunos añadidos relativamente simples al "demonio" LOCKSS, esta mejoría haría la siguiente arquitectura económicamente viable: un demonio LOCKSS funcionando en una máquina local con almacenamiento en un servicio de almacenamiento de objetos.

Resumen elaborado por Natividad Escavias Extremera