

iDA: Preservando el acceso a bases de datos a muy largo plazo

Eladio Gutiérrez¹, Sergio Romero², Oscar Plata³

Resumen— El proyecto iDA (*Immortal Database Access*) tiene como objetivo el ofrecer una visión general de las posibilidades tecnológicas para la preservación de los sistemas informáticos desmantelados (*decommissioned*) y los datos que contienen. Tras el desmantelamiento de un sistema como una base de datos, a menudo es deseable preservar partes del sistema para poder restaurarlas posteriormente en caso necesario.

Basado en la tecnología del proyecto previo iVM (*Immortal Virtual Machine*) se pretende desarrollar herramientas de desmantelamiento que preserven los contenidos de forma permanente en un soporte físico, incluyendo la regeneración de los datos junto con sus funcionalidades de búsqueda.

El objetivo es garantizar la inmunidad a la obsolescencia técnica, protegiendo el acceso y las funcionalidades de búsqueda para futuros usuarios, previniendo la corrupción de datos, y garantizando el acceso a largo plazo que sirva como «copia de seguridad definitiva de último recurso» tanto para las generaciones futuras, como para uso presente (recuperación inesperada, consideraciones legales, etc.).

Para ello se plantea la construcción de un motor ROAE (*Read-Only Access Engine*), que actúe de interfaz hacia los usuarios futuros (o del presente) que deseen recuperar la información desmantelada, almacenada en el sustrato a largo plazo utilizado (cinta filmográfica). Este motor complementa el estándar de almacenamiento elegido SIARD (*Software Independent Archiving of Relational Databases*), que aunque bien captura los datos y metadatos del sistema desmantelado no es capaz de capturar la información clave sobre las funcionalidades de consulta para buscar, procesar y recuperar información cuando la base de datos estaba activa.

Palabras clave— Preservación de información digital, iVM (*Immortal Virtual Machine*), iDA (*Immortal Database Access*), SIARD (*Software Independent Archiving of Relational Databases*), ROAE (*Read-Only Access Engine*).

AGRADECIMIENTOS

El presente trabajo ha sido financiado mediante el proyecto “iDA: Immortal Database Access – Long-Term Recovery and Access to Decommissioned Database Systems“, Eureka Eurostars-3, ref. E!1622, 2022. En el proyecto iDA colaboran las siguientes instituciones: Piql (empresa, Noruega), Tedral (empresa, España), el Centro Noruego de Computación (NR), y la Universidad de Málaga.

REFERENCIAS

- [1] “iDA: Immortal Database Access - long-term recovery and access to decommissioned database systems,” EUREKA Eurostars project, reference number: E!1622, 2022-2024.

¹Dpto. de Arquitectura de Computadores, Universidad de Málaga, e-mail: eladio@uma.es

²Dpto. de Arquitectura de Computadores, Universidad de Málaga, e-mail: sromero@uma.es

³Dpto. de Arquitectura de Computadores, Universidad de Málaga, e-mail: oplata@uma.es

- [2] H. Bruggisser et al., “SIARD format specification 2.0,” Tech. Rep., eCH: e-Government Standards, Suiza, 2015.
- [3] International Organization for Standardization (ISO), “Space data and information transfer systems – Open archival information system (OAIS) – Reference model,” Tech. Rep., ISO, June 2012.
- [4] Ivar Rummelhoff, Eladio Gutiérrez, Thor Kristoffersen, Ole Liabø, Bjarte M Østvold, Oscar Plata, and Sergio Romero, “An abstract machine approach to preserving digital information,” *IEEE Access*, vol. 9, pp. 154914–154932, 2021.
- [5] “Immortal Virtual Machine - solving the problem of file format and infrastructure obsolescence,” EUREKA Eurostars project, reference number: E!12494, 2018-2021.
- [6] Eladio Gutiérrez, Sergio Romero, and Oscar Plata, “Diseño del compilador de la máquina virtual inmortal iVM,” in *Actas de las XXXII Jornadas de Paralelismo*. 2022, Sociedad de Arquitectura y Tecnología de Computadores (SARTECO).
- [7] Jędrzej Sabliński, Alfredo Trujillo, et al., “Piql – Long-term preservation technology study,” *Archeion*, vol. 122, 2021.
- [8] Bjørn H Brudeli, “A holistic approach to digital preservation,” in *SMPTE 2014 Annual Technical Conference Exhibition*, 2014, pp. 1–11.
- [9] Richard M Stallman and the GCC Developer Community, “GNU Compiler Collection internals (for gcc version 10.2.0),” *Free Software Foundation*, 2020.
- [10] Corinna Vinschen and Jeff Johnston, “The Red Hat newlib C library,” 2021, <https://sourceware.org/newlib/> (retrieved on Mar 1, 2021).
- [11] Jeremy Bennett, “Howto porting newlib: A simple guide,” 2010, <https://www.embecosm.com/appnotes/ean9/ean9-howto-newlib-1.0.html> (retrieved on Mar 1, 2021).
- [12] Eladio Gutierrez, Sergio Romero, and Oscar Plata, “ivm64-gcc: The iVM Compiler,” 2024, <https://github.com/immortalvm/ivm-compiler>.
- [13] Eladio Gutierrez, Sergio Romero, and Oscar Plata, “ivm64-as: The iVM Assembler,” 2024, <https://github.com/immortalvm/ivm-compiler/tree/master/gcc-12.2.0-ivm64/gcc/config/ivm64/asm2bin>.
- [14] Eladio Gutierrez, Sergio Romero, and Oscar Plata, “Yet another (fast) iVM emulator,” 2020, <https://github.com/immortalvm/yet-another-ivm-emulator>.
- [15] Ivar Rummelhoff, “Formal specification of VM and I/O devices and description validation (Eurostars Programme, iVM project),” Tech. Rep., Norwegian Computing Center, Oslo, Norway, 2021.
- [16] Thor Kristoffersen, Bjarte M. Østvold, and Ivar Rummelhoff, “Process definition and guidelines for system decommissioning,” Tech. Rep., Norwegian Computing Center, Oslo, Norway, 2023.
- [17] Thor Kristoffersen, Ivar Rummelhoff, and Bjarte M. Østvold, “The DbSpec executable specification language,” Tech. Rep., Norwegian Computing Center, Oslo, Norway, 2023.
- [18] “SQLite home site,” <https://www.sqlite.org/>.
- [19] Eladio Gutierrez, Sergio Romero, and Oscar Plata, “iVM filesystem generator,” 2024, <https://github.com/immortalvm/ivm-fs>.
- [20] “ioctl.tty(2) — Linux manual page,” https://man7.org/linux/man-pages/man2/ioctl_tty.2.html.
- [21] “tinyXML2: a simple, small, efficient, C++ XML parser,” <https://github.com/leethomason/tinyxml2>.
- [22] “zlib: A massively spiffy yet delicately unobtrusive compression library,” <https://www.zlib.net/>.
- [23] “Prepared SQL statement,” https://en.wikipedia.org/wiki/Prepared_statement.
- [24] “DBML - database markup language,” <https://dbml.dbdiagram.io/home/>.

- [25] Andrew Lindley, “Database preservation evaluation report - SIARD vs. CHRONOS,” in *International Conference on Preservation of Digital Objects*, 2013, vol. 10.
- [26] Bruno Alexandre Alves Ferreira, *Database Preservation Toolkit: A relational database conversion and normalization tool*, Ph.D. thesis, Universidade do Minho (Portugal), 2016.