

Modular emulation as a viable preservation strategy

Jeffrey van der Hoeven, Hilde van Wijngaarden

Koninklijke Bibliotheek, the National Library of the Netherlands
The Hague, The Netherlands

Jeffrey.vanderhoeven@kb.nl, Hilde.vanwijngaarden@kb.nl

Abstract. Emulation is the only strategy to ensure long-term access to digital objects in their original environment. The National Library of the Netherlands (KB) and the Nationaal Archief of the Netherlands believe that emulation-based preservation is worth developing and has to be tested. This short paper proposes a new model for emulation called modular emulation that will allow us to develop a working prototype for the rendering of digital objects in the future.

1 Why emulation?

Digital preservation does not end with the careful storage of digital objects. With its e-Depot [1] in place, the Koninklijke Bibliotheek (KB) ensures that digital objects are safely stored for the long term. However, keeping them accessible requires a continuous effort working out strategies for permanent access. The usability of digital objects is threatened by rapid innovations in computer technology. Therefore strategies have to be developed to ensure access to digital objects for the long term. Preservation strategies can be divided into two groups: migration and emulation. The problem with migration is that it often creates small errors and that it is not a one-time event, but has to be applied for all objects periodically with an increasing risk of error-propagation. Moreover, migration may cause loss of functionality. Emulation recreates the original environment in which digital objects can be rendered in their authentic form, without the need for periodic conversions. This way the original object is preserved as well as the functionality offered by the environment. When functionality matters, emulation is the only effective strategy. Existing emulators - created for different purposes - prove that emulation is not too complex and can be cost-effective.

2 Modular emulation

Based on the results of emulation research at the KB and inspired by existing ideas of both Jeff Rothenberg's virtual machine approach [2] and Raymond Lorie's Universal Virtual Computer (UVC) [3], a new model on emulation has been developed: modular emulation. This model (see figure 1.1) focuses on the recreation of the original hardware environment in which digital objects can be executed together with their

system and application software. This way, PDF documents, databases and interactive multi-media applications can be rendered authentically.

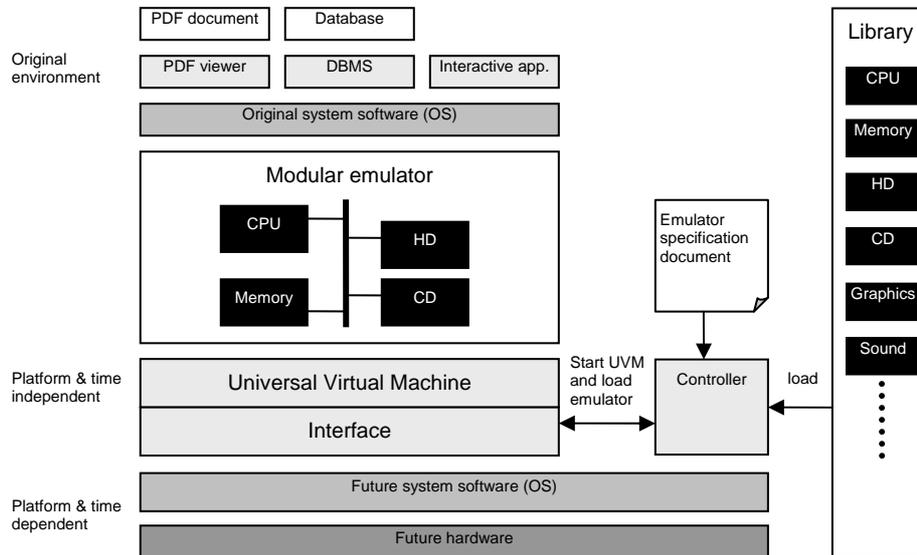


Figure 1.1: Conceptual model of modular emulation

The core of this model is defined by the modular emulator that is able to run the original system and application software. The emulator consists of distinct modules, each of them emulating specific hardware functionalities, like a CPU, memory or storage device. Each module can be reused and rearranged to create different emulators. All modules are preserved in a module library. Based on an emulator specification document that defines which modules should be used, the controller loads the required modules and creates a new emulator on the fly. On top of future hard- and software, a Universal Virtual Machine (UVM) will create the desired platform independency. In turn, the UVM will execute the modular emulator.

In 2005 and 2006, an actual modular emulator will be developed incrementally by the KB and the Nationaal Archief of the Netherlands. This emulator will be platform dependent at first, but will later be extended to run on a UVM.

References

1. Oltmans, E., van Wijngaarden, H.: Digital Preservation in practice: the e-Depot at the Koninklijke Bibliotheek. In: VINE, vol. 34 (1), (2004), 21-26
2. Rothenberg, J.: An Experiment in Using Emulation to Preserve Digital Publications. Koninklijke Bibliotheek, The Hague, The Netherlands, 2000. Available at: <http://www.kb.nl/coop/nedlib/results/emulationpreservationreport.pdf> (accessed 9 June 2005)
3. Lorie, R.A.: Long-term archiving of digital information. IBM Research report, IBM Almaden Research Center, San Jose, Almaden, 2000.