Initial Findings from the 2021 Software Preservation Network Hosted Emulation Services Pilot

Ethan Gates
Software Preservation Analyst, Yale University Library
User Support Lead, EaaSi Program of Work

The Emulation-as-a-Service Infrastructure (EaaSii) grant-funded pilot service work is spending 2021 investigating the feasibility of emulation as a scaled, high-quality, cloud-hosted service using the EaaSii platform.

Exclusive access to a pilot hosted service was offered to organizational members of the Software Preservation Network, with SPN acting as a consortial “customer” and first point of contact while EaaSii oversees deployment, training, support, and data synthesis.

The pilot remains ongoing, but initial findings and feedback have already provided critical insight for future service design and implementing emulation in digital preservation workflows.

Cited use cases for emulation included:
- Design files
- Digital art
- Early text-based/word processor formats
- Data reproducibility
- Research software
- Augmented virtual reality
- Games
- Mobile archaeology/retro-tech research
- Born-digital government records
- Archiving and/or migrating digital objects – of any kind

Emulation services can not be siloed

Consensus grew on the value of coordinated efforts that are not necessarily in-scope for the EaaSii platform itself:
- Defined and shared acquisition policies for collection software (dependencies) to avoid redundant collecting
- Mapping simulation-related metadata to existing schema and standards – e.g., PREMIS, CodeMeta, MODS, RDA
- Enriching related third-party services with regard to treating software as objects of collection and preservation – e.g., Zenodo, Wikidata

To put emulation services into production at their organization, participants reckoned they would have to consult with:
- University/library/department IT
- Digital library services
- Legal/civil
- Collection management services
- Public services, reference desk staff, and/or subject liaisons
- Data services
- Research and scholarly services
- Preservation services
- Metadata/cataloging services

Platforms Testing
- What can the EaaSii platform do?
- What should it do?
- What questions does hosted emulation answer, and which are still open?

Preliminary Google Cloud deployment statistics (July 26 - June 30)

- 1 x 16-core CPU
- 32 potential simultaneous simulation sessions
- 500 GB solid-state storage available
- ~330 unique emulation sessions run
- 65 “Public”, 54 “Private” Emvironments (~60 GB saved)
- 101 digital objects uploaded (~24 GB)

Challenges to offering a hosted emulation service, identified by the EaaSii team:
- Active time zone spread (Australia to U.S. East Coast) prevents clear maintenance windows
- Large majority of User Support Lead and Chief Programmer’s time directed to hosted service debugging, documentation, and service requests; limited availability for other areas of EaaSii program and/or non-hosted EaaSii users
- Balancing requests for new features with infrastructure “invisible” improvements

Contact
ethan.gates@yale.edu
https://forum.eaasi.cloud