

Olive

The Olive Project (CMU, olivearchive.org)

Olive (Open Library of Images for Virtualized Execution)

Observations:

- A strength: use streaming for remote VM execution (called "Internet Suspend/Resume"), so a user can interact with a VM before the entire VM is downloaded, similar to Internet video.
- Also related with Brown Dog's DAP. Only client software is available, server software not released yet. Might be usable for extractors that require a special OS/software environment. Software only available for Linux. Software may not be mature enough. June 2013 presentation listed "thin-client execution in OpenStack" as a future work item.
- Good overview presentation: May 2013, presented to Library of Congress, "Not a Graveyard: The Olive Archive Project". Good technical presentation of the client software VMNetX: June 2013, "[Building VMNetX with qemu and libvirt](#)".

What is it

A collaborative (originally between Carnegie Mellon University (CMU) and IBM Research) project seeking to establish a robust ecosystem for long-term preservation of software, games and other executable content. Born at CMU, Olive addresses the current gap in preservation technology by providing a curated environment for the preservation and distribution of executable content.

-- From olivearchive.org

Executable content examples: simulation models, tutoring systems, expert systems, data visualization tools, company's product information page that are dynamically generated that customizes the content and appearance at run time.

Goals

- Enable libraries to achieve their mission of preserving scholarly records
- Provide a platform for preserving, searching, extending, and distributing executable content
- Establish a workflow for this preservation
- Determine the obstacles to preservation and distribution
- Retain executable content in its original state so that end users may interact with it freely

Team

- Prof. Mahadev Satyanarayanan (Satya), in CS, CMU

Architect of Andrew, Coda files systems

- Gloriana St. Clair

Dean Emerita of University Libraries, CMU

- Erika Linke

Associate Dean of University Libraries, CMU

- Jerome McDonough

Associate Prof, UIUC

- Anita de Waard

VP of Research Data Collaborations, Reed Elsevier

- Daniel Ryan

Curator for Executable Content, CMU

...

Funding

- IMLS National Leadership Grant (Institute of Museum and Library Services, a government agency)

Oct 2012, \$497k over 2 years.

Goal: To understand what types of content can be ingested into Olive and to determine how executable content archive can fit into existing trusted repository standards such as OAIS, OCLC, CRL, JISC.

- Sloan Foundation Grant.

Jan 2013. \$400k over 2 years.

Goal: To develop the technical framework for Olive, to plan for an effective organizational structure to sustain the archive and provide access to executable content.

Test cases

- Great American History Machine

Built by Prof. David Miller at CMU in the 1980's, provides GIS for early census data

- ChemCollective

Award winning educational software built by Prof. David Yaron

- Mystery House

Apple II game for which experimental metadata was generated under IMLS-funded Preserving Virtual Worlds 2

- Doom

Popular PC game with heavy graphics released in 1993

- Information Systems

Special issue of the Elsevier journal which will include executable content

How It Works

- Executable content is preserved in a full VM to ensure long-term compatibility

- Relevant VM archives can be found using content and metadata searches

- VMs are streamed using the ISR technology.

Like Internet videos, users may begin interacting before a full VM is downloaded. Streaming technology offers immediate VM access by sending only the data required for execution.

- Users may save VM changes to their own computers and apply them to subsequent launches.

Software

- Client: VMNetX.

Runs on Linux, using the open-source KVM VM monitor. Can execute VMs directly from any web server; no special server software is required. On GitHub. License: GPL v2. <https://github.com/cmusatyalab/vmnetx>

- Server: not released, plan to release under GPL v2.

Client releases: VMNetX

0.4.4 - bugfix release, 2014-05-31.

0.4.3 2014-03-05.

0.4.2 2013-12-20.

0.4.1 2013-11-07.

Three main ideas

- ISR (Internet Suspend/Resume). A technology to stream VMs over Internet. -- From CMU.

Disk is demand-paged over the network.

- A technique for indexing and searching VM image contents. Searching relies on signature matching, instead of user-specified metadata.

- A technique to incrementally composing a VM image.