

# *OldWeb.today*

*Keeping Retro Web Sites Available on  
the Modern Web (including Flash!)*

*Ilya Kreymer, Webrecorder Lead Developer*



*OLDWEB.TODAY*  
*JS Browser Emulation* **NEW**

The Webrecorder project is focused on advancing open source software development and research in the following key areas:

- FOSS web archiving tools to create and view web archives
- Highest-fidelity capture and replay
- Integrate with existing archival systems
- **Exploring intersection of web archiving and software emulation**
- Empower anyone to create, use and share web archives
- Making web archiving more accessible via decentralized and p2p technologies

# About Webrecorder

*Web archiving for all!*



Webrecorder

*Web archiving for all!*

# About OldWeb.Today

- Browse old web sites with old browsers!
- Original version launched in 2015 with emulators running in Docker containers still available ([classic.oldweb.today](http://classic.oldweb.today)) but requires infrastructure and maintenance (eg. was built using a now-obsolete version of Docker)
- Latest OldWeb.today created in early 2021
- Client-side emulation - all emulators run in the browser itself
- Should improve as WebAssembly and JS emulators improve



Browser window showing the URL: <https://oldweb.today?browser=ns3-mac#19960101/http://geocities.com/>

Page title: **OLDWEB.TODAY**  
*JS Browser Emulation* **NEW**

Browser menu: File Edit View Go Bookmarks Options Directory Window 9:55 PM

Navigation buttons: Back Forward Home Reload Images Open Print Find Stop

Location: <http://geocities.com/>

Browser / EMULATOR: Navigator 3

URL: <http://geocities.com/>

BROWSE LIVE WEB  
 BROWSE ARCHIVES AT:  
 1996-01-01 00:00:00  
 ▶ Web Archive Sources

Status: Running

[How it Works / View Source](#)  
[Classic OldWeb.today](#)  
 A project by: **Webrecorder**

---

TechWire just got more reporters...more news, and of course, it just got a whole lot better. You should come see what all the talk is about. <http://www.techweb.com> (or just click here).

**GEOCITIES** YOUR HOME ON THE WEB

Our communities are home to the most popular collection of **FREE HOME PAGES & E-MAIL** on the web. Please join or visit one of our 29 neighborhoods today.

- **ENTER HERE**
- **INFORMATION**
- **NEIGHBORHOODS**
- **WHAT'S NEW**
- **WHAT'S COOL**
- **WHAT IS GEOCITIES?**

[Free Home Pages & Free Member Email](#)    [Advertiser Information](#)

**Today's Cool Homestead** [GeoCities Daily Audio Update](#)  
[HotSprings 1837](#)  
 So you hit the snooze bar ten times every morning. You might be lazy. But then again, you might have a sleep disorder. Find out here.

**GeoCities News of the Day - 10/22/96**

**GEO+PLUS** Experience GeoCities on a new level! **Subscribe to GeoPlus**

**Introducing GeoPlus!**  
 Use up to 10MB for your site. Earn double GeoPoints. Post live news feeds on your home page. Get a bundle of free software from McAfee Associates. And more. [Start today](#) with our pre-launch subscription offer!

**Is modern society getting you down?**  
 Come meet Giogg, the Internet's only Neanderthal advice columnist, this week on the

# Web Archiving

## *What is web archiving?*

“Web archiving is the process of collecting portions of the World Wide Web, preserving the collections in an archival format, and then serving the archives for access and use.” [International Internet Preservation Consortium](#)



Webrecorder

*Web archiving for all!*

# Web Archives

**Most well known:**

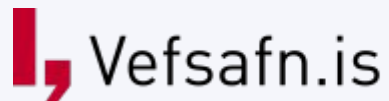
INTERNET ARCHIVE  
**WayBackMachine**

# Public Web Archives

**Most well known:**



**But also many others out there:**



*Many institutions run their own wayback machines...*



**Webrecorder**

*Web archiving for all!*

# Accessing Public Web Archives

Many web archives make the original web content (without banner, any changes available via) a de-facto convention, using the “id\_” modifier: `https://<archive-path-prefix>/<timestamp>id_/<url>`

Examples:

- `https://arquivo.pt/wayback/19961226182558id_/http://www.geocities.com/`
- `https://web.archive.org/web/19961013213730id_/http://www.yahoo.com/`
- `http://wayback.vefsafn.is/wayback//19980210175524id_/http://www.geocities.com/`

Given a web archive that supports this, possible to load archived web pages dynamically given a URL and Timestamp.

(Note: Still need a Cross-Origin-Resource-Sharing (CORS) proxy - as most web archives do not support CORS)



Webrecorder

*Web archiving for all!*



## Emulators used in Oldweb.today

- **v86** - popular x86 Emulator written in JS, and WebAssembly  
(Used for Window and Linux Environments)  
<https://copy.sh/v86/>
- **Basilisk II** - 68k Mac Emulation, then ported to JS and WebAssembly  
(Used for MacOS Environments)  
<https://jamesfriend.com.au/basilisk-ii-classic-mac-emulator-in-the-browser>
- **Ruffle** - Flash-only emulator written in Rust, running WebAssembly  
<https://ruffle.rs/>



# Supported Environments

- Each environment was prepared in v86 or Basilisk II
- For Basilisk II, disk images created using desktop version
- For v86, disk images exported via web version after installing browser.  
(also restoring emulator state for faster loading)

		NCSA Mosaic 2	
		NCSA Mosaic 3	
		MacLynx 2	
		Navigator 3	SHOCKWAVE 3 JAVA 1.0
		Navigator 4	SHOCKWAVE 3
		IE 4	SHOCKWAVE 3 JAVA 1.0
		Navigator 4	FLASH 8 JAVA 1.1.5
		IE 5	FLASH 9 JAVA 5.0
		IE 6	FLASH 9 JAVA 5.0
		Firefox 10	FLASH 32
		Opera 12	FLASH 32
		Ruffle (Flash)	FLASH

# JS/WebAssembly Emulators

Great Support for:



Video



Audio



Networking



# JS/WebAssembly Emulators

Great Support for:



Video



Audio



Networking

**Need a JS/WASM based TCP/IP stack to support networking!**



Webrecorder

*Web archiving for all!*

## How to connect emulators to web archives?

- Web Archives need to be accessed via HTTP, over TCP/IP
- V86 supports sending Ethernet frames over a WebSocket
- Basilisk II port in JS had no networking support, but Basilisk II does, and can get Ethernet frames.
  
- Want to avoid relying on external infrastructure as much as possible.

**Ethernet Packet — ??? — HTTP Web Archive**



# How to connect emulators to web archives? TCP Stack in the Browser

- **Solution:** Extending on excellent work from eaas-proxy:  
<https://gitlab.com/emulation-as-a-service/eaas-proxy/-/blob/master/webnetwork.js>
- PicoTCP compiled via Emscripten to WASM + wrappers for sockets over native JS web streams (full TCP stack)
- Can terminate the HTTP connection from emulator in the users' browser!



# Custom Networking Setup

**Emulator Web Worker:** Run emulator WASM (v86 or Basilisk)

**PicoTCP Worker:** Run PicoTCP Stack

Emulator Worker sends Ethernet frames to PicoTCP web worker over a broadcast channel, which processes them as TCP packet, and passes them to HTTP Server, which then proxies to the live web / web archive...  
... and then sends the response back!



# Connect emulators to web archives

## Specifying URL and Timestamp Dynamically

How to dynamically set the URL and Timestamp in the emulated environment?

Solution: Browser in emulator uses an HTTP proxy pointing to, *10.0.2.2:6082*, redirects to the actual URL

When timestamp is changed, user is asked to reload to get new home page.

URL:

BROWSE LIVE WEB

BROWSE ARCHIVES AT:

► Web Archive Sources

Date Updated!  
Click the *Refresh* button or load a new page in the emulated browser to start browsing at the new date.

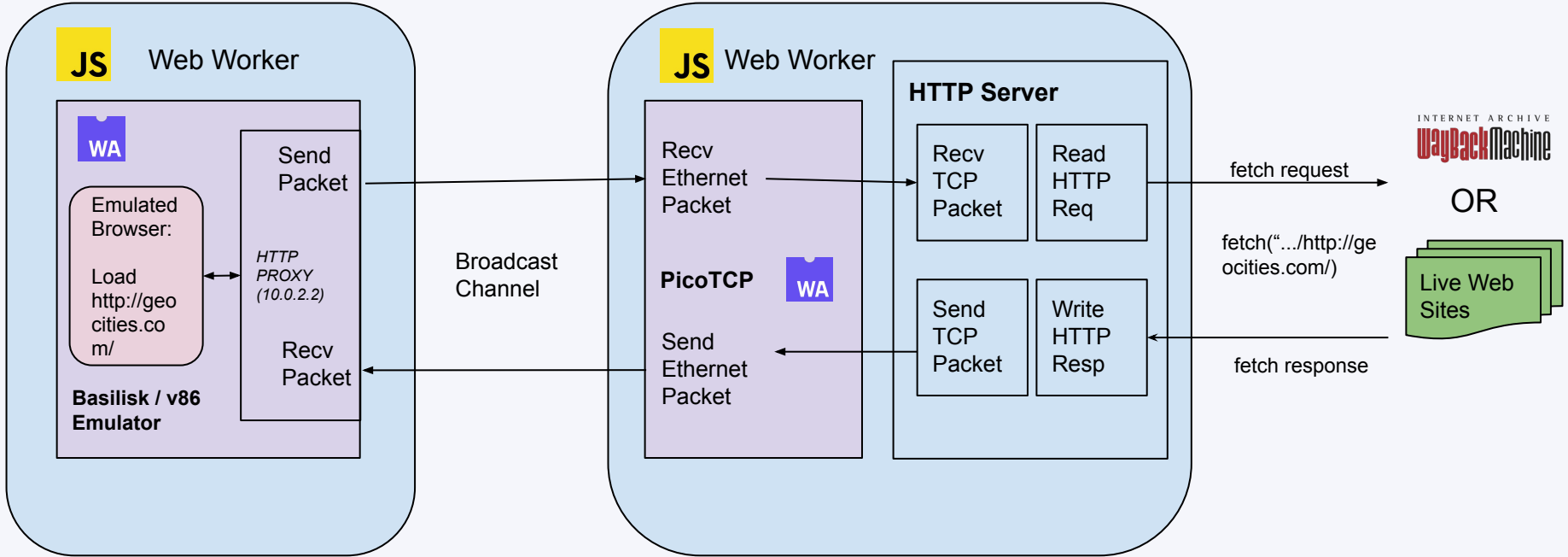


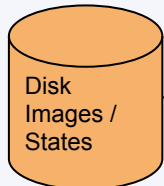
Webrecorder

Web archiving for all!



# OldWeb.today Networking Flow





Disk Images / States



Live Web Sites

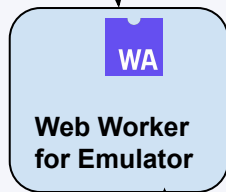


INTERNET ARCHIVE  
<https://web.archive.org/web/1996/http://geocities.com/>

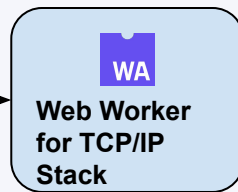


CORS Proxy

*In Browser*



Web Worker for Emulator



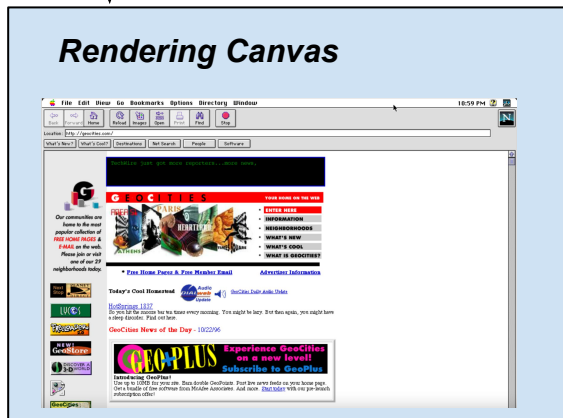
Web Worker for TCP/IP Stack

Oldweb.today Page

URL

Date

Archive Prefix / Live Toggle



Rendering Canvas

**OldWeb.Today Architecture**



Webrecorder

Web archiving for all!

# What about Flash (and Java?)

- ✓ Prepared environments include browsers with various versions of Flash and Java installed!
- ✓ Many Flash projects do work!
- ✓ Can support latest Flash (32)
- ? But only older browsers (upto Firefox 10 / Opera 12)
- ? Many Flash/Java not fully archived in the 90s / early 00s (.swf, .jar files missing)
- ? Perf: Loading can be slow for Flash (and Java)

		NCSA Mosaic 2	
		NCSA Mosaic 3	
		MacLynx 2	
		Navigator 3	SHOCKWAVE 3 JAVA 1.0
		Navigator 4	SHOCKWAVE 3
		IE 4	SHOCKWAVE 3 JAVA 1.0
		Navigator 4	FLASH 8 JAVA 1.1.5
		IE 5	FLASH 9 JAVA 5.0
		IE 6	FLASH 9 JAVA 5.0
		Firefox 10	FLASH 32
		Opera 12	FLASH 32
		Ruffle (Flash)	FLASH

# Ruffle Emulator

- ✓ Also supported in oldweb.today!
- ✓ Web page loaded directly in native browser, emulator only loaded for Flash content
- ✓ Much faster than a full-OS emulation (no need for custom networking)
- ? Not all of Flash is supported, yet (no AS3)
- ? Re-implementation of Flash, may be subtle differences.





## Other Approaches to Supporting Flash in Web Archives

	<i>Supports All Flash (include AS3)?</i>	<i>Runs at full fidelity? (Video + Audio + Speed)</i>	<i>No Infrastructure CPU Cost Per User</i>	<i>Available in Tools:</i>
Ruffle	Not Yet	Yes	Yes	<b>oldweb.today</b> archiveweb.page replayweb.page
OldWeb.today JS Emulators	Yes	Not Yet	Yes	<b>oldweb.today</b>
Remote Browsers	Yes	Yes	No	pywb-remote-browsers conifer classic.oldweb.today



## Links: oldweb.today

- Blog Post about oldweb.today:  
<https://webrecorder.net/2020/12/23/new-oldweb-today.html>
- Main Repo:  
<https://github.com/oldweb-today/oldweb-today>
- Basilisk II Fork:  
<https://github.com/oldweb-today/macemu>
- V86 Fork:  
<https://github.com/oldweb-today/v86>
- EaaS JS Network Based + PicoTCP Emscripten Build:  
<https://gitlab.com/emulation-as-a-service/eaas-proxy>

## Other Webrecorder Tools to help with Flash

- Emulators running server-side in Docker containers:  
<https://github.com/webrecorder/pywb-remote-browsers>
- ArchiveWeb.page Extension - allows archiving Flash via Ruffle emulator in the browser  
<https://archiveweb.page> (injects Ruffle)



**Webrecorder**

*Web archiving for all!*

# Future Work for oldweb.today

- Optimization!
- Adding more emulation environments to support newer browsers + newer OSs
- Make networking more robust / debug occasional errors
- Rewrite https URLs for older browsers
- Support for adding different besides Internet Archive
- Ability to upload an archive directly + create archive while browsing (integration with other Webrecorder tools)





# Webrecorder

*Web archiving for all!*

**THANK YOU**



[info@webrecorder.net](mailto:info@webrecorder.net) (Email)



[webrecorder](https://github.com/webrecorder) (GitHub)



[Subscribe](#) (RSS)



[webrecorder](https://www.youtube.com/webrecorder) (YouTube)



[webrecorder\\_io](https://twitter.com/webrecorder_io) (Twitter)



[Discuss](#) (Forum)

<https://webrecorder.net/>



Webrecorder

*Web archiving for all!*