

# 模擬保存之發展 ——數位羊皮紙與 EMiL EaaS

The Development of Emulation Preservation: Digital Vellum and EMiL EaaS Project

楊耀名 Yang, Yao-Ming

英福達科技股份有限公司實驗室經理

Laboratory Manager, InfoDoc Technology Corporation

## 壹、前言

近二十年來科技不斷的發展，在此高度數位化的時代，人們越來越仰賴電子檔案，但由於電子文件保存困難，如何完整保存及典藏愈趨重要。你是否也有過這樣的經驗，當以為自己完整備份郵件時，想開啟 5 年前的重要資料卻發現無法開啟檔案，諸如此類的問題會越來越常發生。因為現在大部分的文字、影像已大量數位化，如果這些資料無法於未來讀取，那麼 21 世紀的歷史及文化，只有極少數文獻資料會流傳到後世，少數專家已逐漸開始正視此問題，尋求解決的方案，也有專家建議，珍貴的文件、照片等資料，還是印出以紙本保存，也便於後代保留及儲存。

由於網際網路的發展，政府、公司及金融機構也逐漸地將業務電子化，如電子化政府、電子詢價系統等。現今常用的電子郵件、部落格、SNS (Social Networking Service) 張貼的文字，許多都是以數位形式儲存在硬碟或是雲端上的資料。當數位資料越來越多的同時，如何完整保存是非常重要的課題。

網際網路之父 Vinton Cerf (2016) 提出「數位黑暗時代」的警告，表示由於資訊科技的快速變遷，軟、硬體設備不斷更新而造成檔案格式很快過時，未來可能面臨資料因軟硬體設備過時而無法讀取，以至於難以繼續保存而消失，形成數位黑暗。

因此在研究如何保存電子檔案的同時，亦冀望能一併保存其種類及存取資料時所需操作方法與相關訊息。卡內基美隆大學「數位羊皮紙」(digital vellum) 計畫，即為達成此一目標而產生。另外德國佛萊堡大學提出 EMiL EaaS 計畫，提出使用模擬技術提供過時作業系統，使得能在作業系統裡操作過時檔案以達成保存數位資料之目的。

## 貳、電子檔案長期保存的種類

目前發展出系統保存、轉置、模擬及封裝等策略來進行資料保存。歐陽崇榮 (2008) 認為軟硬體及設備可使用系統保存的方式，透過保存電腦系統設備，以達成電子檔案可繼續使用，但每臺設備保存是非常耗成本及空間的方法。另外過

時格式可採用轉置的方法，但由於轉置後格式亦存在資料損失的風險，未來能否被讀取且完整使用是一個未知的風險。模擬則可使得資料於新電腦環境中操作，且模擬概念則是操作原始環境及應用軟體，使得檔案可以開啟且保存完整的檔案資訊，有鑑於此，本文著重介紹模擬保存的技術與方法。

何謂模擬？簡單來說就是在現有的電腦架構下，透過模擬器來模擬原來老舊或過時的電腦系統之硬體或軟體。藉由模擬的行為來呈現當時的數位化物件，透過這樣的方式可以讓使用者體驗到該物件最真實的外觀或感覺。於是，模擬成為解決電子檔案長期保存的策略之一。

## 參、數位羊皮紙介紹

### 一、源起

所謂羊皮紙是古代用來記錄文字、訊息的書寫材質，使得歷史得以保存下來。而今檔案資料大量的電子化、軟硬體設備不斷的更新，這些數位檔案如何妥善保存及於未來使用避免面臨數位黑暗風險。Vinton Cerf (2016) 認為不能只靠單純的保存數位位元資料，而要保存的是整個數位資料的意義，於是參與卡內基美隆大學的「數位羊皮紙」(digital vellum) 計畫。該計畫乃是要建立一個可成功及正確性的去讀取老舊的檔案並可互動的環境，亦可解讀為一個數位的科技生態系統。簡單來說，即在保存電子檔案時，亦需留存詮釋資料，內容包含檔案之作業環境，如讀取設備、作業系統、應用軟體、版本等，以確保檔案資料得以被讀取。透過這樣的保存機制，即可以完整實現透過軟硬體環境來模擬過時設備，經由安裝作業系統及應用系統，使所保存的數位資料得以呈現或編修。且為避免版權爭議，需在特殊法律架構允許下來實作。主要目標亦在永久保存

學術記錄、檔案資料，提供平台保留檔案資料的原始狀態並且可於未來執行檔案內容，達成解決數位資料無法讀取之問題。Vinton Cerf 及卡內基美隆大學團隊後來透過 Olive (Open Library of Images for Virtualized Execution) 專案之技術架構解決了數位羊皮紙的關鍵問題。

### 二、系統運作方式

數位羊皮紙在 Olive 專案中採網路串流虛擬機系統，即在一臺虛擬機上同時模擬多臺的硬體設備且模擬許多還在使用或已經不存在的操作機器。由於不同年代發展出的設備，所運用的控制指令不同，如未精確處理，其呈現的結果和當時會有差異。模擬最重要是須確保設備在執行上的保真度。在模擬該硬體設備時，首先要撰寫該設備的模擬軟體，且精確地描述硬體運作方式。

此架構應用中使用快照的技術，將硬體設備、作業系統及應用程式以映像檔方式保存，如需重建當時環境時，則將映像檔載入即可。結構如圖 1。

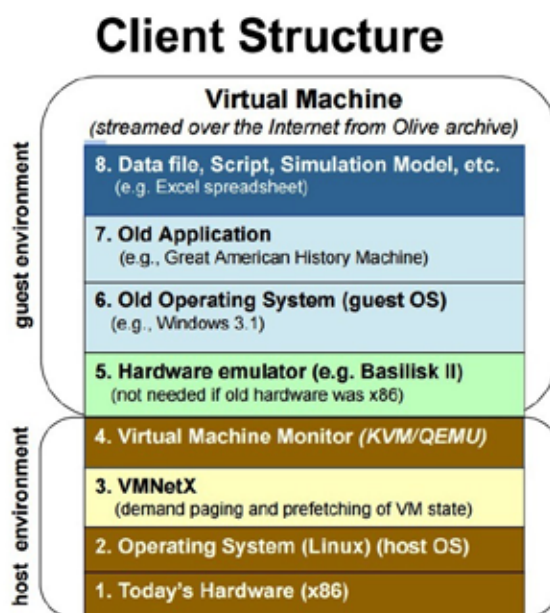


圖 1. Olive 專案環境架構

資料來源：Mahadev Satyanarayanan, Gloriana St. Clair, Benjamin Gilbert, Yoshihisa Abe, Jan Harkes, Dan Ryan, Erika Linke, Keith Webster (2015) One-Click Time Travel. Retrieved from <https://olivearchive.org/documents/>

為能在現行設備運作執行模擬程式，通常需要很大的記憶體空間。另為使虛擬環境被廣為應用，首先需要將其映像檔的詮釋資料和內容放在網路上，俾利搜尋及下載使用。由於映像檔很大，下載時間很久，可能會降低使用意願，因此，該團隊採用類似視訊檔的串流技術，讓使用者在虛擬機器下載當下即可使用。另外，映像檔也可以客制化，下載保存在個人電腦中，或重新安裝其他軟體後，上傳至平台上。另為加速作業效率，平台將依使用者瀏覽歷程及執行紀錄，預測所需要的環境予以推薦，使用者可免除逐一挑選設定，如圖 2。

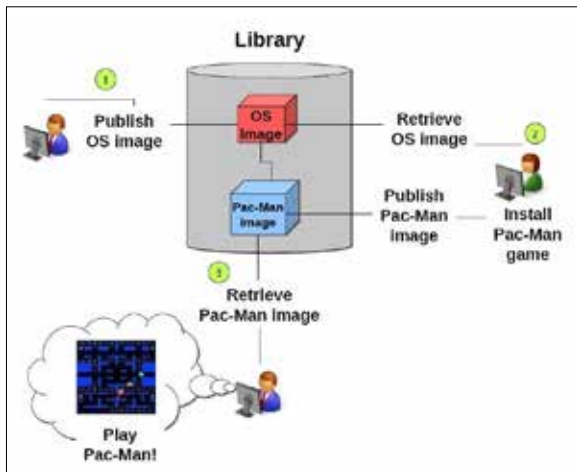


圖 2. Olive 執行情形

資料來源：Olive Archive. Retrieved from <https://olivearchive.org/about/>

Olive 專案之虛擬機器映像檔 (VM Image Representation) 包含機器描述 (Domain XML)、磁碟映像檔 (Disk Image) 及記憶體映像檔 (Memory Image) 3 部分，如圖 3 所示。其中機器描述，包含如模擬器類型、驅動程式名稱、匯流排等相關資訊。如此，全部封裝在一個可以隨時放進虛擬機器裡的檔案，以達到運作之目的。

圖 4 範例即是用 QEMU 模擬在微軟作業系統 Win3.1 執行應用程式，如系統保真度很高，將使這模擬看起來與原來環境執行結果一模一樣。

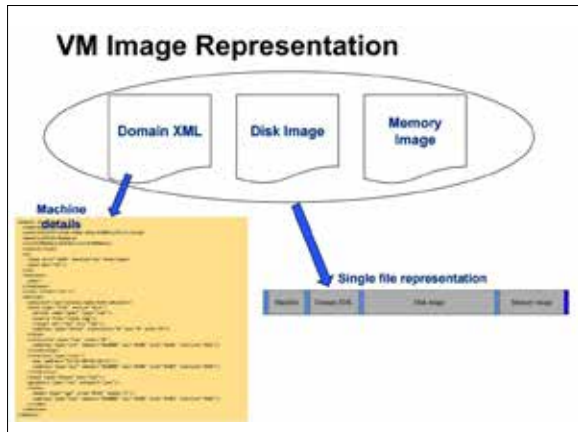


圖 3. VM 映像檔

資料來源：Vint Cerf. (2016) Digital Vellum and Archives. Retrieved from <https://www.nitrd.gov/>

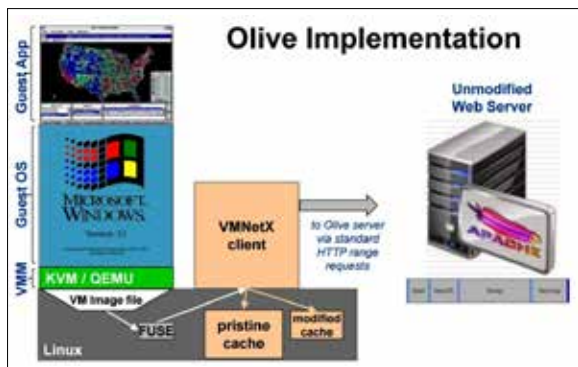


圖 4. Olive 模擬實例

資料來源：Mahadev Satyanarayanan et al. (2015) One-Click Time Travel. Retrieved from <https://olivearchive.org/documents/>

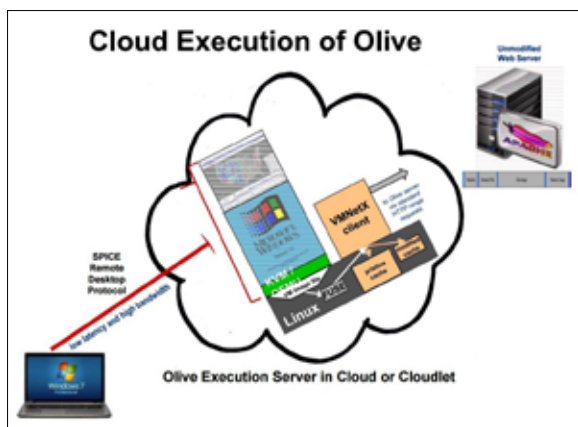


圖 5. Olive 雲端平台運作

資料來源：Vint Cerf. (2016) Digital Vellum and Archives. Retrieved from <https://www.nitrd.gov/>

Olive 專案採用了雲端平台的運作提供模擬服務，將是一個解決數位資料流失的方法。採用這樣的運作機制，即可以遠端來執行，並可有效解決數位檔案無法讀取問題。如圖 5。





圖 6. 操作 Microsoft Office 6.0 for Windows 3.1 on x86 畫面  
資料來源：Mahadev Satyanarayanan et al. (2015) One-Click Time Travel. Retrieved from <https://olivearchive.org/documents/>

### 三、成果及挑戰

2011 年該團隊已成功展示如何在虛擬機內執行 Windows 3.1 作業系統如圖 6，和早期 Mac 等舊系統；以及這些虛擬映像檔如何儲存在中心程式庫中，並透過 Internet 進行串流傳輸。為使這個計畫持續運作並不簡單，其中一個就是規模問題，你必須思考決定需要保留什麼，什麼時候該保留，要保留多久，還有保留它們的條件是什麼。另一個最大的挑戰是必須持續提升處理能力、系統硬體效能及精確性。此外，由於多樣化的資料格式，必須搭配不同的軟體程式，在有限的資源裡，很難決定哪些軟體及版本是需要保存放置於映像檔中。因此，在電子檔案長期保存上，適當的壓縮檔案格式，並採取開放軟體，方得以維持電子檔案長期可讀取性。

## 肆、德國計畫 EMiL EaaS 介紹

### 一、源起

本節介紹德國佛萊堡大學團隊提出以模擬的方式進行檔案資料的保存。德國佛萊堡大學於 EMiL EaaS 專案中提到大部分的情況下呈現數

位資料最佳的方式為使用其當初操作的系統及應用軟體，因為這些系統包含數位資料的重要資訊及屬性，因此能提供使用者真實的交互使用體驗。另外也提到數位資料是非常容易被改變的，且隨著使用的過程，系統也容易受到改變。為了確保數位檔案永久保存並且能於未來被使用，提出了 EaaS (Emulation as a Service)，該專案目標為提供一個數位資料的原始環境，從而保持其原有的相關資訊及實用性。

### 二、運作方式

模擬為數位保存的策略，使用模擬技術複製大量數位資料的方法，未來將成為機構接受數位資料保存的方式，因此完整的模擬服務需求將不斷成長。為了提供使用者良好的模擬服務，需要在網際網路或基礎設施上運行可擴展的分佈式系統模型。

到目前為止，模擬已被視相關技術專家專屬領域。此外，由於繁瑣的準備和技術設置程序，模擬無法良好的擴展。因此，該團隊發展了一個可擴充的模擬服務模式：模擬即服務 (Emulation as a Service)。

模擬即服務 (EMiL EaaS) 架構簡化了對保存數位資料的存取，允許終端使用者於不同模擬器運行原始環境。現成的模擬元件提供彈性的 Web 服務 API，允許開發個性化和定制的數位保存流程。

EMiL EaaS 的基本架構是抽象的模擬元件，用於標準化部署和隱藏單一模擬器的複雜性。每個模擬元件將特定的模擬器類型封裝為具有統一的應用程式介面 (API) 集合的抽象元件。這樣，不同類的模擬器變得可相互操作、互換或相互連結。此外，各模擬元件可以有效地建置在大規模節點或雲端中，如圖 7。

EMiL EaaS 目前有過去主流和當今作業系

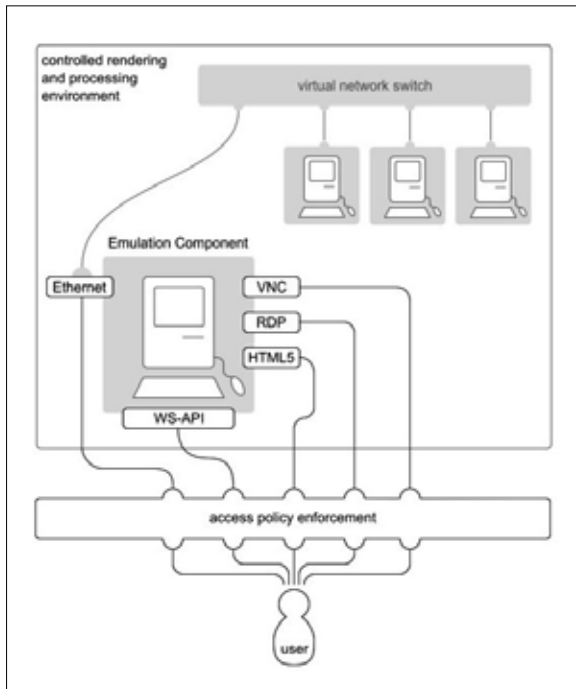


圖 7. EMiL EaaS 運作架構圖

資料來源：bwFLA -Emulation as a Service. Retrieved from <http://bw-fla.uni-freiburg.de/eaas.html>

統的模擬組件，例如 PPC，m68k，Intel 的 x86 等，以及主要的作業系統，例如，OS/2、MS Windows、Mac OS 7 和更高版本等可透過模擬呈現。

為了可擴展及具有成本效益，模擬元件需要僅在需要時部署。為此，需要適用於硬體和軟體部署的框架。目前 EMiL EaaS 支援 MAAS/Juju 以及 OpenStack 技術架構，如圖 8。

EMiL EaaS 提供使用者試用帳號及密碼，並提供一個簡單目錄清單提供，讓使用者選擇欲使用的作業軟體，如圖 9。

通過選擇目標對象，將啟動 EMiL 存取工作流程。網頁目錄系統將使用者 ID 傳遞到 EMiL EaaS 框架。接著 EMiL 啟動模擬系統，使用 HTML5 技術在瀏覽器中呈現作業系統。經由使用者可透過操作作業系統，開啟過時軟體以及過時檔案。圖 10 為使用 EMiL 模擬操作 MAC OS 9 畫面。

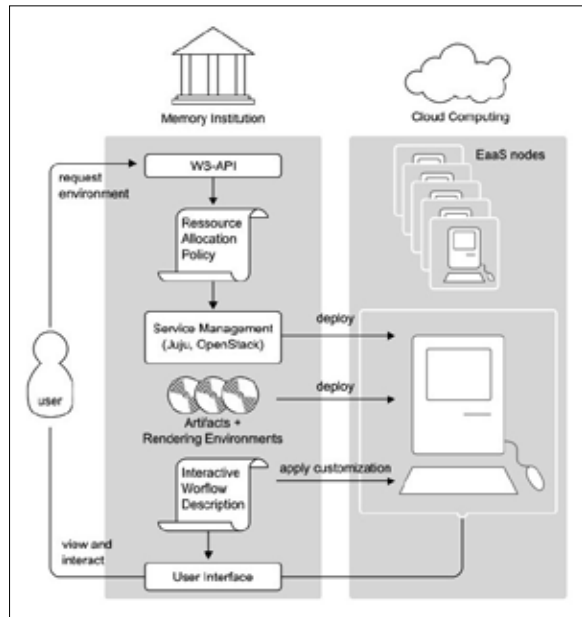


圖 8. EMiL EaaS 運作架構圖

資料來源：bwFLA -Emulation as a Service. Retrieved from <http://bw-fla.uni-freiburg.de/eaas.html>

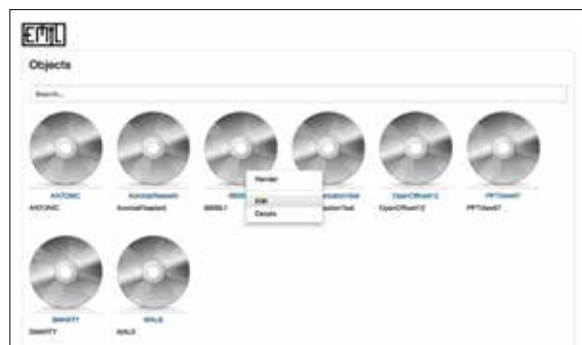


圖 9. EMiL EaaS 模擬清單

資料來源：Emulation as a Service. Retrieved from <http://demo.bw-fla.uni-freiburg.de/>



圖 10. 透過 EMiL EaaS 操作 MAS OS 9 畫面

資料來源：Emulation as a Service. Retrieved from <http://bw-fla.uni-freiburg.de/publications.html>

### 三、成果及挑戰

EMiL EaaS 提供了模擬服務的策略，於此 EMiL EaaS 提供系統即為服務的觀點，並使用相同的技術存取現今及過時的系統。將複雜及機密的檔案或文件保存在儲存機構中或由第三方提供者處理，如此一來可確保機密檔案的安全性，並且使用者不需要安裝任何軟體。另外 EMiL EaaS 透過使用雲端技術以助於標準化及改進服務，使第三方用戶（如檔案典藏機構）可以接受。在分佈式 EMiL EaaS 模型中，檔案數位化的成本可以共享。通過相互專業化，EMiL EaaS 可以包含特定領域的需求。模擬和模擬器將變得更容易處理。目前該專案已成功運行過時檔案及軟體於 OS/2，MS Windows，Mac OS 7 等早期作業系統。此外解決未來使用者設備及技術不斷變化為該專案面臨最大的挑戰，為確保順利於系統運行過時檔案，需針對特定的雲端後端平台進行長期維護。

對於有長期保存電子檔案需求的機構，創建 EaaS 是一個良好的方法。如果機構標準化歸檔且確保能於其網絡元件的長期可用性，則不會產生額外的歸檔成本。模擬即服務 (EaaS)，保存計畫及保存成本是固定的，僅由模擬器和模擬系統的數量決定。目前，EMiL EaaS 支持 9 個不同的模擬器，能夠模擬 15 個不同的傳統作業系統平台。每個模擬元件可用於透過網頁服務連結在各種歸檔工作流程中使用，以達成電子檔案長期保存。

### 伍、比較分析

若就我國運用電子檔案保存使用角度來分析，國內目前對於雲端運作架構技術已經漸趨成熟，不論是參考數位羊皮紙或 EmiL EaaS 的專案運作方式皆可朝向雲端方式建立服務並提供服

務方向來進行研究與開發。目前我國各機關在系統建置上大量使用 VM 的技術架構來建構相關的應用系統，可規劃於該 VM 應用系統下架後之保存應用，如任務型或特殊網站於任務完成後，倘若未規劃後續轉檔或移轉機制，可以採用模擬運作方式，將相關 VM 進行保存，在網路頻寬足夠及人力或財政允許下適度提供保存應用，亦不失為一種檔案保存方式，更可提供後續學術面或教育面使用，讓使用者能了解當時的數位資訊。

數位羊皮紙及 EMiL EaaS 專案皆提供雲端化平台進行模擬操作，模擬操作方面由於 EMiL EaaS 較早進行研究，因此作業系統方面 EMiL EaaS 支援作業系統較數位羊皮紙多。另外兩計畫之比較分析表及主要特點可參考表 1。

### 陸、結論與建議

21 世紀為數位化的時代，當科技不斷發展，越來越多數位化服務、資料存取越便捷的同時，也相對面臨由於科技進步導致一些過時檔案軟體無法使用的問題，未來也許會發生含有重要資訊的過時檔案面臨相關作業硬體或是作業系統過時而無法開啟檔案的狀況。

處理數位檔案保存在各國愈發重視下逐漸成為不容忽視的課題，也因此激盪出各相關專家學者開始探討及著手找出解決問題的方法。有鑑於此，國際上有越來越多相關領域的專家、學者開始重視數位資料保存的問題，進行各種研究。數位羊皮紙及 EMiL EaaS 計畫皆提出以模擬技術解決長期保存數位檔案的問題。該等計畫持續運作並不簡單，因此，如何維持電子檔案長期可讀取性，建議可以幾個面向思考：

#### 一、政策引導面

如同檔案分類是檔案整編及保管之基礎工

作，有利於後續檔案管理、查檢與應用，電子檔案保存可依此精神進行分析歸納並建立具保存價值之分類，如訂定典藏價值網站、特殊應用系統等，擬定相關之保存蒐集或分類標準，使機關能夠遵循。

## 二、發展統整面

以相關分類之機關屬性選出適當機關作為統籌機關，以避免各機關各自為政導致整體資源分散，無法有效豐富整體保存面架構內容。

## 三、服務應用面

透過積極發展雲端服務方式並強化雲端服務安全性，如能夠在排除頻寬問題等狀況下，進而提供 end user 應用之服務，使得相關保存資料在

完善保存下，減少數位黑暗之情形，方可達到能提供保存資料之應用。

另外，數位羊皮紙及 EMiL EaaS 計畫之模擬技術皆面臨到相關作業系統及應用系統的版權問題，建議相關政府機關訂定專案法規以為因應系統版權問題。此外在模擬操作使用上偶爾會出現系統反應較慢的情形，希望未來串流相關技術突破後能提供更快的連線效率，以增加系統穩定性。此外 EMiL EaaS 部份操作頁面為德語，如果能於未來提供多國語言畫面，將更易於提供使用者操作。專案的進行及研究需大量的人力及經費與多年持續的研究才能開枝散葉，除了研究學者積極參與外，整合民間及相關政府機構投入經費，共同推動研究計劃，才能實現電子檔案長期保存之願景。

表 1. 兩計畫之比較分析表

	數位羊皮紙	EMiL EaaS
研究國家	美國	德國
研究年代	2011~ 至今	2005~2014
運作架構	Client 安裝、Cloud	Cloud
可模擬作業系統	Android 4.1、Debian、Fedora、Red Hat、Ubuntu、Windows 7	Red Hat Linux、Windows Serial、MS DOS 6.20、Apple Mac OS、Atari 1040ST、BBC Microcomputer、Commodore 64、IBM PC-5150、Macintosh System、Ubuntu
應用面	透過使用者端安裝 VMNetX 軟體才能連結。惟該 Server Site 軟體尚未公開發布。	透過公共雲方式（Google Compute），發佈三款使用 CDROM 的老舊電腦遊戲線上模擬 <a href="http://www.theverge.com/2015/4/17/8436439/theresa-duncan-chop-suey-cd-rom-preservation">http://www.theverge.com/2015/4/17/8436439/theresa-duncan-chop-suey-cd-rom-preservation</a>
主要特點	<ul style="list-style-type: none"> <li>使用串流技術，僅發送執行所需的數據來提供 VM 虛擬機作即時存取用</li> <li>可以單獨執行或導入單一文件，而無需重新建構 VM 虛擬機</li> </ul>	<ul style="list-style-type: none"> <li>模擬運用 web services 方式提供服務（不須軟體的安裝）</li> <li>透過遠端進行資料運算</li> <li>建立可管理的模擬工具</li> <li>選擇模擬環境後自動安裝</li> <li>可以與現有的歸檔解決方案整合</li> </ul>

資料來源：作者自行整理。

---

## 參考文獻

1. 歐陽崇榮 (2008)。數位資訊保存策略。台北市：文華。
2. Vinton Cerf (2016) . *Digital Vellum and Archives*. Paper presented at FASTER CoP.
3. Cody Brooks (2015). A digital dark age may be coming, Google VP warns. Retrieved from <http://www.digitaltrends.com/photography/digital-dark-age-strike-21st-century-history-internet-creator-warns/> (February 19,2015)
4. Mahadev Satyanarayanan, Gloriana St. Clair, Benjamin Gilbert, Yoshihisa Abe, Jan Harkes, Dan Ryan, Erika Linke, Keith Webster (2015). One-Click Time Travel. Retrieved from <https://olivearchive.org/documents/>
5. Dirk von Suchodoletz, Klaus Rechert and Isgandar Valizada (2013). Towards Emulation-as-a-Service: Cloud Services for VersatileDigital Object Access. *The International Journal of Digital Curation*, 8(1), doi:10.2218/ijdc.v8i1.250
6. Mahadev Satyanarayanan, Vasanth Bala, Gloriana St. Clair, Erika Linke (2011, October). *Collaborating with Executable Content Across Space and Time*. Proceedings of the 7th International Conference on Collaborative Computing.
7. Networking, Applications and Worksharing, Orlando, FL. Retrieved from [https://liber2014.univie.ac.at/fileadmin/user\\_upload/DOEVL\\_events/Kongressfiles/2014-05-20\\_LIBERworkshop\\_emulation\\_en\\_kb.pdf](https://liber2014.univie.ac.at/fileadmin/user_upload/DOEVL_events/Kongressfiles/2014-05-20_LIBERworkshop_emulation_en_kb.pdf) (Jeffrey van der Hoeven in 3rd LIBER workshop, Vienna, May 20th, 2014)