IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.12417

Web OS for Document Management

Vignesh Bala R[1], J.Bhuvana[2]

Student, MCA, Jain deemed to be university, Bangalore, India[1]

HOD, School of CS & IT, Jain deemed to be University, Bangalore, India[2]

Abstract: Web OS for document management is a software designed for managing documents, providing a secure vault to store critical business documents. Companies now prefer storing important files in Document Management systems instead of file servers. Web OS for document management functions like a regular file system, but with enhanced capabilities. Instead of storing files on a local hard disk, files are stored within the system, facilitating easy access for individuals who need them. It operates like a network file server but with more sophisticated features.

Keywords: document management, Web, OS, file server, important files.

I. INTRODUCTION

Managing content has become a laborious task for organizations due to the sheer volume of content produced on a daily basis. IT professionals invest significant amounts of time in managing content and ensuring that it adds value to the company. To address the challenges associated with traditional content management techniques, Web OS for document management has been introduced. This tool aims to ease the burden on IT professionals, allowing them to be more productive.

All content kinds can be managed with the help of the Web OS for document management. Even if the content is spread across a number of servers and physical storage devices in a networked environment, it is kept in a secure repository that is subject to compliance requirements and is housed in a single environment. It offers superior security compared to a file system. With Web OS, access to a file can be restricted to specific users or groups, preventing unauthorized access, modification, or deletion.

The content in an organization can originate from various sources, such as CRMs or ERPs. Thus, Web OS for document management can be regarded as a comprehensive solution for managing all types of digital content used in an organization. Its use can also facilitate a more efficient decision-making process.

Web OS for Document Management consists two main components. These components make up of the whole system:

i. **Proxy server:** This server is a temporary storage unit. The files get stored here as cache files, but it will soon move to another destination where there is no other path to connect to except this server. This acts as user interface to the users and contains user credentials to for checking the authorization, permissions and certificates to perform the kind of activities that the respective users have. Then the files are brought back from the Main Server and viewed here as this is a UI to the user. Then after saving the changed file is sent to the Main Server. This is the server that is connected to the network and has access to all devices that logs in.

ii. **Main Server:** This is the primary server where all kinds of files are saved. The primary and the only purpose of this server is the storage of files and media. Here files are retrieved and saved. The main advantage of such an architecture is security and compliance. It also has high scalability, elasticity, efficiency and accessibility. The other reason is that users can access their files from their local device from anywhere using their unique URL and password from any browser

II. LITERATURE REVIEW

In this essay, web-based e-content management is discussed. In particular, this article discusses the technologies, standards, problems, and difficulties surrounding content repository (the methods for saving and accessing content) and content contribution (the methods for producing content). for controlling and managing digital content and processes), workflow (content management), automation services, and automation services lifecycle. It gives a rundown of the products that are now available and explains how they can be applied to the deployment of digital libraries. The last ten years, and particularly the previous five, have seen advancements in the development of digital libraries. Using the internet, libraries may quickly and easily make text, speech, and visual content (such as animation, graphs, pictures, and videos) available to consumers. the intended result.

Managing the constantly rising volumes of online material on the WWW has become crucial in terms of repository, retrieval, workflow, and distribution. Because to developing technologies, services can now be delivered anytime and everywhere. Although libraries, museums, and archives are working to leverage technology to conserve and distribute significant content for their customers, a lack of funding and industry standards have forced them to look for additional

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023
DOI: 10.17148/IJARCCE.2023.12417

collaboration options and methods. The National Science, Mathematics, Engineering and Technology Education Digital Library, the Digital Library Federation (DLF), and the Library of Congress's National Digital Library Program (NDLP) are a few such. The primary purpose of the DLF is to "bring together digitised content from around the nation and overseas that will be available to researchers, academics, and students. [1]

According to a University of California at Berkeley study, the amount of digital records that must be kept across different information domains will grow at an unprecedented rate in the coming years. Regulatory requirements, such as the Sarbanes-Oxley Act, policy requirements, such as the need for governments to preserve records of their public policies and services, and history-focused requirements, such as the preservation of cultural heritage for future generations, all pose archiving and preservation challenges for organisations. Hence, mass storage design presents the initial difficulty for archiving system designers. Yet, it just solves a fraction of the more significant problem, which is digital preservation. The topic has received discussion on the internet over time. A ground-breaking examination of preservation strategies is provided by Thibodeau with their advantages and disadvantages. There is still disagreement over the right course of action. Nonetheless, the community has started using the following methods: migration, emulation, preservation format, and Universal Virtual Computer. [2]

Regardless of the technology and software that was used to create the data, electronic records must be kept and made available through electronic archive systems meant for long-term storage. In order to store, maintain, and provide search capability and access to electronic documents over a long period of time, the large-scale systems must meet the criteria given below briefly and explored in more detail. An enormous number of digital documents need to be kept across several information domains. According to studies like the "How Much Information" by the University of California at Berkeley, exabyte ranges of data are predicted to become the new standard. Repositories with a petabyte capacity are now in use. [3]

The Xen and VMware-popularized hypervisors are swiftly eroding into commonplace technology. They are suitable for many usage scenarios, but some call for system virtualization with high levels of efficiency and isolation. Examples include hosting facilities, PlanetLab, the Grid, and HPC clusters. We offer a better solution than hypervisors for these kinds of situations. The method is a fusion of earlier work on general-purpose, timeshared operating systems with resource containers and security containers. Some container-based systems include LinuxVServer, Virtuozzo for Linux, and Solaris 10. This article presents the design and implementation of Linux-VServer as an illustration of container-based systems. Also, it compares the architecture of Linux-VServer with the most recent Xen generations.[4]

In order to accommodate numerous diverse applications operating concurrently on server systems, data centres are increasingly using virtualization and consolidation. The design and evaluation of server platforms, however, is still focused on the performance modelling of a single, highly parallel application or a collection of homogeneous workloads running concurrently. Since server virtualization is anticipated to be used by the majority of future datacenters, this study examines the difficulties associated with simulating virtual machine (VM) performance on a datacenter server. Based on the most recent multi-core servers and the server virtualization benchmark vConsolidate, we demonstrate that the VM modelling challenge necessitates solving three major issues: (a) By simulating the use of visible resources (such as cores, memory, and I/O devices), (b) simulating the competition for intangible resources (such as shared microarchitecture resources, cache, and memory bandwidth); and (c) simulating the implementation costs of virtual machine monitors (or hypervisors). By defining a VM performance modelling approach and conducting a thorough case study based on the vConsolidate benchmark, we take a first step towards solving this issue. We end by listing unresolved issues for future effort. [5]

Cloud computing has gained popularity as a result of the expanding usage of high-speed Internet technology. Users work using Web-based software and storage rather than local software and storage while using cloud computing. These browser-based apps provide a desktop-like user interface and behaviour. [6]

The Web is transitioning from a mechanism for distributing documents to an application framework. The desktop, along with the programmes and merchants that have made their fortunes there, might be completely overthrown by the new Web. Of course, not everyone believes that evolution and revolution are synonymous. The WebOS is criticised as being more hot air than hot news. But, proponents of WebOS are confident that they understand all that is going on and that it has already started. The Vision of WebOS Neither the totality of WebOS nor a single thing exists (yet). The Vision of WebOS Neither the totality of WebOS nor a single thing exists (yet) (yet). A notion called WebOS lays out a roadmap—or perhaps a gauntlet—for how the convergence of software and the Web can upend how we compute and challenge long-standing market dynamics. One-box solutions—a physical container carrying all the software and hardware you need to execute your applications—have dominated the history of personal computing thus far. A desktop operating system, such as Apple's OS X or Microsoft Windows, offers a number of interfaces between the hardware within the box and the applications you use—your word processors, graphic design programmes, web browser, etc. [7]

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.12417

A software system must actually be developed and maintained by a large team of individuals. The Integrated Development Environment (IDE), the primary tool for programmers, is a solo tool that aids in understanding and system modification by a single programmer. Such an IDE does not take use of any design and implementation skills that other team members may have. We suggest that this issue be solved by testing out novel approaches to maximise developer cooperation and deriving information from prior IDE interactions. Our strategy, known as ADINDA, is based on turning the IDE into a collection of interconnected services that are available through a web browser and enhanced with Web 2.0 tools. These services will assist developers with typical IDE activities as well as the informal communication and teamwork requirements of software development projects. We discuss our goal, strategy, obstacles faced while constructing ADINDA, and preliminary findings in this document. [8]

Today's web browsers are intricate, have sizable trusted computing bases, and provide hackers simple access to cutting-edge computer systems. The Illinois Browser Operating System (IBOS), a new operating system and a new browser that decreases the trusted computing base for web browsers, is introduced in this study. We eliminate practically all traditional OS components and services from our trusted computing base by directly mapping browser abstractions to hardware abstractions in our design, which exposes browser-level abstractions at the lowest software layer. We demonstrate that this architecture is adaptable enough to handle new browser security rules, maintain compatibility with conventional apps, and significantly reduce browsing overhead [9].

In this work, we introduce Gazelle, a multi-principal OS-based safe web browser. The operating system that controls resource distribution and protection only for web site proprietors is the kernel of the Gazelle browser. This structure reveals complex design flaws, such as cross-protection-domain display and events protection, that no prior effort has been able to identify. We go into further detail about these problems and offer thorough remedies. It is feasible to convert a current browser into a multi-principal OS that produces much higher security and robustness with tolerable speed, according to our prototype implementation and assessment experience. [10]

III. OBJECTIVE

✧ The cloud storage technologies has become the widely used system in many businesses and startups. Everything is now saved and accessed via cloud storages, since it has become easily accessible and stored. AWS, Azure, Citrix, Nessie, Sharepoint and many other companies offer outsourcing to businesses and earn profits. They provide high scalability, security, compliance and data indexing and develop new methodologies to stay in the market.
✧ While cloud storage has become a popular option for data storage, it also has some potential disadvantages, including:
✧ Dependence on Internet Connectivity: Access to cloud storage is dependent on a stable and fast internet connection. If the internet connection is slow or unavailable, users may not be able to access their data.
✧ Security Risks: Storing sensitive or confidential data in the cloud can pose a risk if the cloud service provider's security measures are inadequate. Data breaches can occur, and if they do, it can expose sensitive information to unauthorized users.
✧ Potential Data Loss: While cloud storage providers take measures to ensure data is safe, it is still possible for data to be lost due to technical issues or system failures. This can result in the permanent loss of important files.
✧ Limited Control Over Data: When storing data in the cloud, users have limited control over where the data is physically stored or who can access it. This may be a concern for users who require strict control over their data.
✧ Cost: While many cloud storage providers offer free or low-cost plans, storing large amounts of data in the cloud can become expensive. Users may also be required to pay for additional features or services.
✧ Dependence on the Provider: When using cloud storage, users are dependent on the cloud service provider's reliability and uptime. If the provider experiences technical issues or goes out of business, users may lose access to their data.
✧ So, Web OS Document Management is the best alternate solution to overcome all these drawbacks and continue serve large businesses and government data more securely, more complacent, better data control and meta data control, cost efficient and much more advantages over cloud storage method.

IV. PROPOSED METHODOLOGY

The proposed system varies from the previous version by two main issues:
i. The documents are only downloaded to the system without the option to view it online. This occupies unnecessary space in the local device and in turn fills up secondary storage devices and causes the users to clear up old data, sometimes leading to accidentally deleting important files.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023

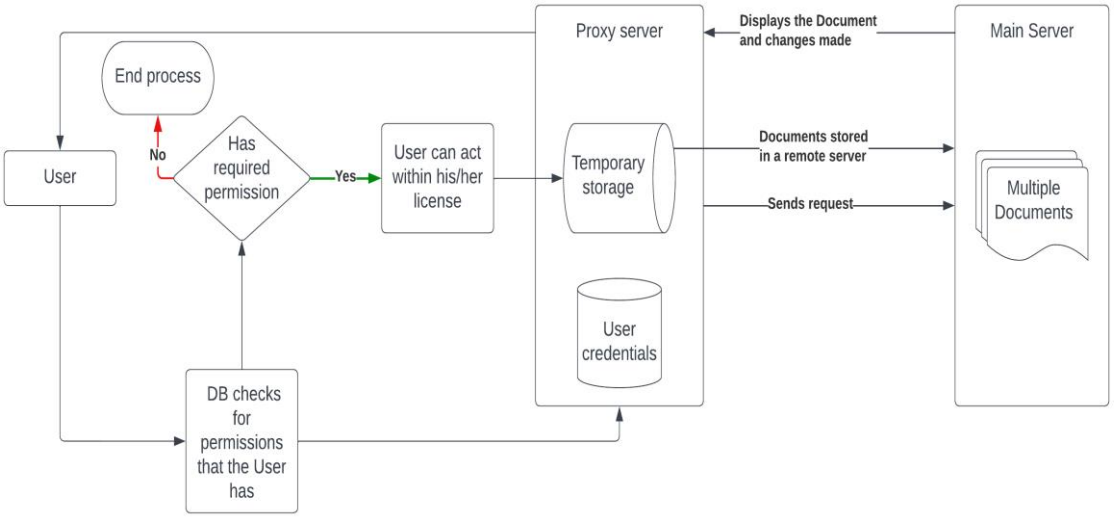
DOI: 10.17148/IJARCCE.2023.12417

ii. The user can view old changes to the document because it always saves the one previous version of the current document saved in the Main Server. It helps in keeping tab of the changes made and can be reverted if the changes were wrong and retrieve the old data. Saving the previous version can help view the history of changes made in organizations. For example, employee data always change and updating the sheet is important to find who is responsible for the corresponding role. Saving the history of such data which are changed, can always help in finding who was in the company/organization that was responsible for that role at that time or when they left the company/organization.

DM is a content management system that provides a centralized repository for storing, managing, and sharing documents and other digital assets. Here's a high-level overview of how it works:

- DM uses a database to store all content-related information, including documents, metadata, and user permissions.
- The DM Main Server is the core component of the system. It handles content retrieval and storage, version control, security, and workflow management.
- Content can be uploaded to DM using several methods, such as through the web interface, via email, or using a desktop client. When a document is uploaded, it is stored in the repository and assigned a unique identifier.
- DM allows users to define custom metadata fields, which can be used to organize and classify content. Users can also create virtual folders, which are views of content based on specific metadata criteria.
- DM provides a flexible security model that allows administrators to control access to content based on user or group permissions. Permissions can be set at the folder or document level.
- Workflow management is another key feature of DM. Workflows define a series of tasks or steps that a document must go through before it can be approved or published. Workflows can be customized to fit specific business processes.
- DM provides integration with other enterprise applications, such as CRM or ERP systems. This integration allows users to access content from within other applications, and to store content related to those applications within DM.

Overall, DM provides a comprehensive solution for managing content throughout its life-cycle, from creation to archiving.



V. CONCLUSION

In conclusion, Web OS for document management's impact on business operations will be significant, improving productivity, reducing costs, and increasing revenue. By streamlining document management processes, improving compliance and regulatory requirements, improving collaboration and knowledge sharing, and reducing costs, Web OS for document management has become an essential tool for many businesses. As businesses continue to generate and manage more content, the need for a comprehensive ECM system like Web OS for document management will continue to grow. The software has more access to information, improved customer service, and enhanced collaboration, leading to increased innovation and better quality products. As a result, consumer behavior has changed, with consumers now having more information and better access to products and services. Overall, Web OS for document management has been a critical tool in transforming how organizations manage and share content, leading to a better consumer experience.

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.12417

REFERENCES

[1] "Content management and the future of academic libraries" by Yuhfen Diana Wu and Menxiong Liu.
[2] "Content server system architecture for providing differentiated levels of service in a digital preservation cloud" by QL Nguyen, A Lake
[3] "Content server architecture pattern for evolvability and scalability" by QL Nguyen, A Lake
[4] "Container-based operating system virtualization: a scalable, high-performance alternative to hypervisors" by Stephen Soltesz, Herbert Pötzl, Marc E. Fiuczynski, Andy Bavier, Larry Peterson .
[5] "Modeling virtual machine performance: challenges and approaches" by Omesh Tickoo, Ravi Iyer, Ramesh Illikkal, and Don Newell
[6] "Moving the OS to the Web" by George Lawton
[7] "WebOS: say goodbye to desktop applications" by Aaron Weiss
[8] "Adinda: A Knowledgeable, Browser-Based IDE" by Arie van Deursen, Arie van Deursen, Bas Cornelissen, Andy Zaidman, Martin Pinzger, Anja Guzz.
[9] "Trust and Protection in the Illinois Browser Operating System" by Shuo Tang, Haohui Mai, Samuel T. King
[10]"The Multi-Principal OS construction of the Gazelle Web browser" by Helen J. Wang, Chris Grier, Alexander Moshchuk, Samuel T. King, Piali Choudhury, Herman Venter

