

White Paper



Approach and Benefits of Embarking Rich Media Content Management Systems on Cloud

Nidhi Tiwari

Abstract

News and media companies require dynamic and collaborative systems to deliver up-to-date rich media contents, using innovative means to keep a tap on competition and cost. However due to limited hardware, their systems face challenges in performing optimally for rapidly modulating workloads and data volumes. These can be effectively tackled using cloud computing as it provides performance, scalability and reliability. This paper presents an iterative approach for putting Rich media content management systems on cloud. It also highlights multi-dimensional benefits thus achieved including business agility through system extensibility to quickly offer new services and adopt latest technologies. Oracle being one of the technology providers in cloud, example to implement the proposed architecture using Oracle stack is shown here.

Introduction

News and media companies require a dynamic web presence for better customer attraction and longer engagement. The rich media content management systems provide solution for all of these requirements. For building one of their key assets, quality contents, they offer collaborative platforms for citizen journalism and large/distributed editorial teams through Web 2.0. To make customer revisit their site, personalization and contextualization is provided on their web portal. These systems also integrate with related websites, build channels, and enable B2B interaction to expand business. For business progress analysis and future growth strategies extensive reporting and analytics are generated using their Enterprise Content Management systems.

However, these rich media systems face challenges in meeting these requirements optimally and continuously due to ever changing/growing content, huge size of rich media content, varying user load, etc. Some key culprits for such performance, availability and reliability issues are – time and cost to quickly scale the infrastructure with increasing load; and the maintenance cost of increasing IT systems.

Now the cloud computing with its elasticity provides a cost effective solution for effectively handling these issues. So rich media content systems when run on cloud can result in better IT returns by increasing operational efficiency, resources utilizations, availability and performance with scalable infrastructure.

As these systems in various organizations have evolved over time and would closely drive the way these organizations work, moving the huge systems in one time may hinder some functionalities and also hit lot of road blocks. Therefore here we present an iterative approach to adopt the cloud computing for Rich Media organizations. We also show that using this kind of cloud environment how quickly system can be amended to offer new services and adopt upcoming technologies using a sample service and an emerging technology.

Iterative approach to move Rich Media Content Management Systems on Cloud

With so much of news being published every day, millions of documents would get created in a short duration. And archiving such volume over years will definitely pose lot of cost and maintenance issues. This drives to the point that content is the most important as well as the most bulky part of the Rich Media Systems. Managing the content most carefully and efficiently would hold the key for the success of these systems. The cloud adoption solution's Iterations 1 and 2, focus on these two aspects respectively.

Iteration 1: Rich Media Content using Cloud Storage

As the cloud can provide virtually unlimited storage capacity, moving the huge volume of data in the Content Management Systems will provide the required storage capacity and capability to quickly increasing the capacity as and when required. Figure 1 below shows, the architecture for putting the Rich media content in cloud for infinite capacity requirement.

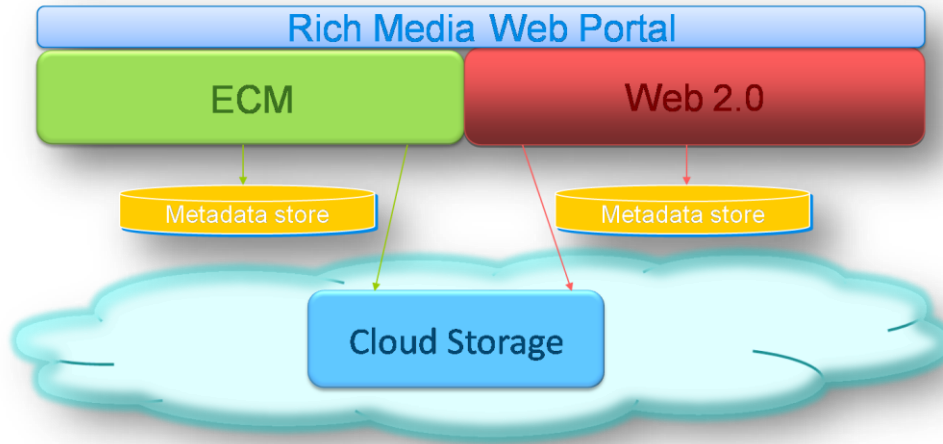


Figure [1]: RMCMS Logical architecture with CMS using Cloud Storage in Iteration 1

This solution suggests to move the content files from local physical storage systems to cloud's storage. The overall Rich Media systems architecture remains same as the existing, except for the file storage now being in the cloud. The content/files are stored on to cloud storage systems like S3 from Amazon and metadata is stored in local database. The identifiers stored in the metadata store are used to write, read and delete the content in cloud storage systems. CMS can use the available API(s) for accessing the content in cloud storage, using the metadata from the local databases. The Cloud storage systems like S3 store the uploaded content as chunks across multiple storage nodes ensuring that even if multiple nodes fail, the content can still be recovered thus offering high reliability.

In cases where the Content management system uses Lucene algorithm for faster content search, then Lucene or similar indexes can be stored on local memory or disk depending on requirement and their size, along with a permanent backup on the cloud storage.

For the security purposes, the access control mechanisms provided by the cloud storage provider can be used while adding/accessing the files from the cloud. The privacy of the content can be assured, by using the encryption/decryption techniques while uploading and downloading the content from cloud storage.

Apart from the storage scalability, reliability and availability, this solution also provides cost benefit, as cloud offers storage at much lower costs, for example S3 costs only 0.10 USD per GB per month.

A good example of a Media site using cloud storage is New York Times' – "TimesMachine". "TimesMachine" is the newspaper's archives site by NYT, which includes full-page image scans of papers back from 1800s. The huge volume content of this is partly hosted on Amazon cloud [1] using the Amazon's storage systems.

Iteration 2: Faster on-demand Processing of Huge storage volume

In second iteration, analytics and content processing can be moved onto cloud to utilize on-demand compute and parallel processing capabilities.

Amazon offers EC2 on-demand computing that can be used to process the content stored in S3. To utilize this elastic infrastructure, Rich media applications' compute intensive tasks like analytics which require to continuously scale-up for processing ever growing data can be moved to the cloud.

Another high performance solution offered by cloud is MapReduce framework, which distributes the processing across multiple nodes and then consolidates the results from each node to produce the final result. Thus it reduces the overall processing time for the job. Such a feature can be used for faster search, reporting and analysis of the content on cloud.

The solution below shows how distributed file system and map-reduce framework can be incorporated into the cloud storage solution.

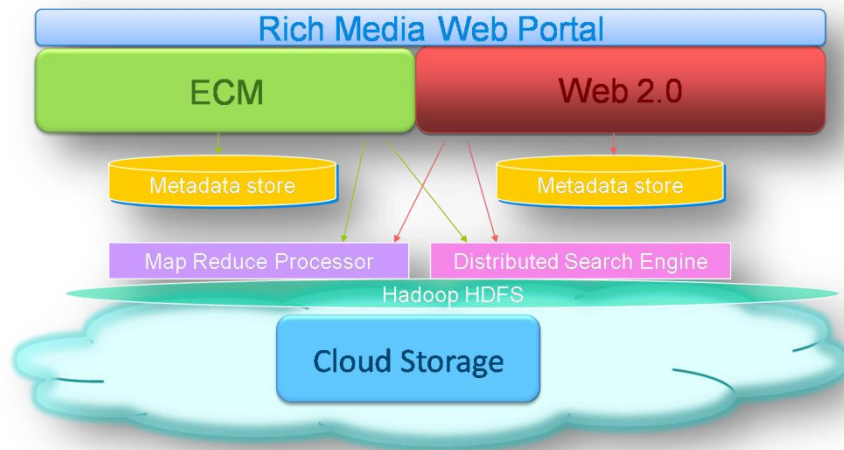


Figure [2]: RCMCMS Logical architecture with CMS and analytical processing on cloud in iteration 2

Iteration 3: Rich Media Content System on Cloud

Once storage and processing concerns for the continuously growing content is tackled, focus should shift to the highly variable aspect of system i.e. workload. Though the ideal case for a media company would be to have regular customers, but with kind of options available and varied interests of people in different kinds of events/news, it doesn't happen so. The load on system keeps varying, depending on the inflow of interesting news on the website etc. Say, if one fine hour the Bulls start ruling the stock market, lot of prospective buyers would log in the website to see news updates on the Stock market. Or if there are any crises being reported in some part of the world, people across the globe would be hitting the site in large number to know more about it. Such spikes occur rarely, and if the organization is not ready to handle those, it will lose the business and subsequent prospective customers. For supporting such large load organization would need to have many fold infrastructures, which would not be utilized at other times, thus impact their return of investment on IT.

Solution for handling such varying traffic would be to have the flexibility of adding/removing the hardware as and when required. Also minimizing the maintenance cost and maximizing the returns from IT. Cloud computing can provide such elasticity, as customers can vary the computing power dynamically depending on their requirement.

Another bottleneck for the faster delivery of huge size media content worldwide is the network. Even if backend is processing all the requests coming to it quickly, the response gets delayed as it has to travel large distances over limited bandwidth. The content delivery network services offered by cloud vendors provide the solution for this problem. The network of servers maintained by cloud vendors at key locations around the globe help distributing content to multiple locations, ensuring faster access. One instance of a media industry getting benefited by use of CDN in cloud is of MSNBC.com. They were able to deliver live video to over 100,000 unique simultaneous users in real time using The Limelight Networks platform

So Iteration 3 proposes the solution with web portal, the ECM and the web 2.0 running in cloud computing environment for higher scalability. The solution shown below in figure 3 has the whole rich media system running on cloud, using the CDN providing high performance.

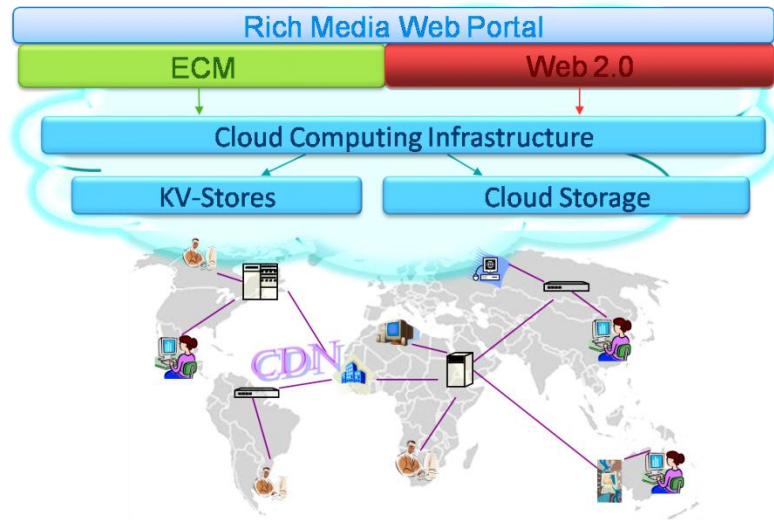


Figure [3]: RMCMS Logical architecture completely on cloud in iteration 3

As the content in the Media systems is document based, and the web 2.0 data is document based, the database can be moved to document oriented Key value stores [4]. These stores will further provide the automated partitioning, replication of data for better availability, scalability and reliability.

When the whole system is in cloud [5], all the components like web portal, content management system and storage are in the same environment, thus improves the performance over the iteration 1 and 2 solution.

The cloud infrastructure consists of distributed hardware, and the distributed computing system provides the fail-over mechanism for uninterrupted services. So moving the Media content management solution to Cloud computing infrastructure improves the availability and reliability of the system.

A good example of rich media website scaling almost instantly using cloud is Animoto. Animoto was a new web 2.0 based website for video creation, upload and presentation, which saw exponential growth in its customer base. To support such high load, it needed to scale up to an astounding 5,000 servers from 40 virtual servers in just four days. And it was able to do so efficiently using cloud saving time and money both. Now the cloud's elasticity allows Animoto to scale dynamically keeping cost low when traffic is low, while still allowing the system to scale up to meet demand when traffic is high.

Iteration 4: Adding New Rich Media Service Exchange @ Cloud

With the availability of Rich Media systems on a scalable environment, lot of business expansion opportunities can be explored. This vast environment can provide a basis to create a collaborative ecosystem for progressing in multiple directions like vendor management, developing partnerships, social engagements, user personalization and so on.

As the complete Rich Media application is available on cloud, tacking a Service Exchange platform on it will evolve it efficiently and effectively to share services and information. Being on cloud this would exhibit scalable and reliable characteristics [6] to provide fast delivery services for multiple departments as well as B2B and B2C offerings. Thus help organization save and earn money.

Having service exchange will help a news industry offer/consume innovative rich media services:

- Delivering news/rich media on mobile/e-mail to individuals,
- Exchange of news with other news agencies,
- Offer analytics to business/government organizations,
- Live business updates to/from organizations/market,
- Fast rich media streaming to multiple vendor/clients, etc.

While offering such services on cloud, organization can benefit by easily providing flexible commercial options like pay-as-you-go, individual transaction cost, service based pricing models, time based price model and so on. Likewise, consuming service on cloud will help reduce time-to-market for new services through optimal cost provider.

Within the organization, a service exchange will help leveraging existing services and data to quickly implement Enterprise SOA for achieving business agility, IT consolidation and quality of service [7].

Oracle Implementation Stack

Oracle has significant offerings for cloud computing. Oracle in partnership with Amazon Web Services (AWS) has made available its products for deployment in public cloud [11] and offers private cloud stack on its grid computing capabilities [18]. So for companies having Oracle technology implementation, a simple and quick way to move their existing systems onto public or private cloud is illustrated here.

Oracle Universal Content Management (UCM) being a complete CMS and web 2.0 ready is a good candidate for the base Content Management system in the implementation. Moreover as explained by David Roe in his blog [12] the readiness of UCM to have its content on Amazon S3 to use virtually infinite storage capacity, increases its applicability in the proposed solution. For web 2.0 component realization, Oracle web center along with Oracle UCM can be used [13]. Next the Oracle WebLogic server image available for Amazon EC2 can be used to implement the rich media portal and application server layer in cloud. While utilizing the distributed cloud environment for faster and scalable request processing, having Oracle coherence [16] with WebLogic server to enable distributed caching will provide the added advantage of quicker data retrieval. Using the Oracle SOA suite on Amazon EC2 [14], a scalable service exchange @ cloud can be implemented. The figure 4 below shows a sample Oracle implementation stack for Rich media systems on Amazon cloud.

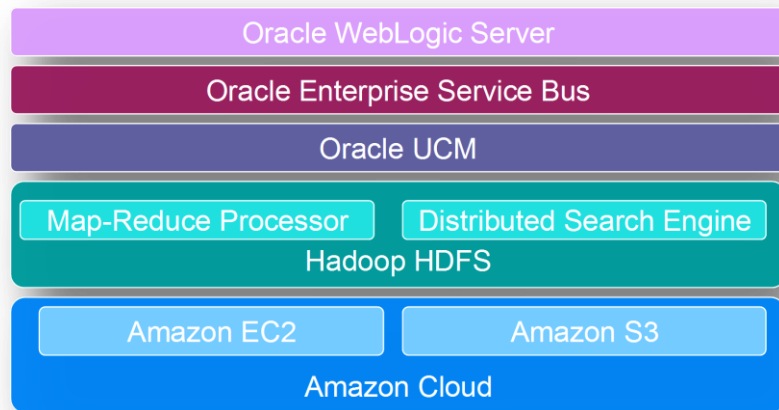


Figure [4]: RMCMS technology architecture on public cloud

As suggested in [17], for companies having data centers optimizing the usage of the current hardware using private cloud could be a natural direction to follow. Media industries with elaborate data centers can also opt for an organic adaptation of their existing systems into cloud, by using Oracle technologies. One alternative for this could be to use Oracle’s grid computing suite [15] beneath the software stack discussed above. Few other sample oracle stacks suggested in [17] and [18] can be customized as per the organizations requirements and features required to be incorporated for media industry as proposed here.

Adopting New Technology: WMSNs based Services

One of the multiple future opportunities that can be explored quickly and cost effectively using cloud environment setup would be of Multimedia Sensor Events as a Service. This is extension of the proposed SEaaS by authors in [10] that can be made available by using the Wireless Multimedia Sensor Networks (WMSNs). WMSNs are networks of wirelessly interconnected devices that can retrieve multimedia content such as video and audio streams, still images, and scalar sensor data from the environment [8]. It is suggested by researchers that by improving the capabilities of current wireless sensor networks WMSNs will boost existing application and lead to several novel applications like: a) multimedia surveillance sensor networks to communicate, process and store data relevant from miniature video cameras installed in various key locations; b) traffic control systems to monitor car traffic and accordingly offer routing advices; etc. [9]. Now as proposed in [10] various WMSNs deployed for such different applications can be brought under one roof as single virtual WMSN through cloud computing infrastructure. It can be exploited by News Company to offer live coverage on distinct topics, reducing the person decencies and delays to B2B as well as B2C services. Further the virtually unlimited cloud infrastructure available can be used to store and process the huge amount of data generated from vast sea of sensors. Thus integrating WMSNs with cloud will make it easy to share and analyze real sensor data or event as a service over internet.

Summary and Conclusion

Rich Media industry systems have the key requirement of managing huge content files and storing them for longer time. In addition, they need to enable the collaboration and personalization, B2B integration, supporting the workload variation with high performance to the end-customer. For successfully overcoming the challenges posed by above requirement, iterative cloud adoption is proposed here. In first iteration as the storage is moved to cloud environment, virtually infinite storage is available for large size and ever increasing media and personalization content. In next iteration, by introducing the HDFS, Map-reduce processor and distributed search engine, QoS requirements of performance, availability and reliability are achieved. Finally, in iteration 3, when complete web system is hosted in the cloud environment; it becomes highly performing, scalable, fault tolerant and available.

Additionally this paper describes how to set-up multi-media, news services exchange on the cloud for generating more business with available infrastructure and data. A futuristic perspective of providing Multi-media Sensor Event as a Service (MSEaaS) over cloud is introduced for News Media industry. They can collect and offer media news on fly to their customers and business partners using the available environment. As the organization doesn't need to invest much in buying the new hardware, cost and risks involved in experimentation is less. Thus the solution comes at Low Cost, no CapEx worries, with little OpEx and Low Risk for new services and ideas implementation.

Acknowledgment

I would like to thank Shyam Kumar Doddavula for his motivation and guidance to complete this work.

References:

- [1] http://www.seattlepi.com/business/375501_cloudcomputing19.html
- [2] MapReduce: Simplified Data Processing on Large Clusters, Jeffrey Dean and Sanjay Ghemawat, OSDI'04: Sixth Symposium on Operating System Design and Implementation, San Francisco, CA, December, 2004.
- [3] <http://hadoop.apache.org/>
- [4] Dynamo: Amazon's Highly Available Key-value Store, Giuseppe DeCandia, Deniz Hastorun, Madan Jampani, Gunavardhan Kakulapati, Avinash Lakshman, Alex Pilchin, Swaminathan Sivasubramanian, Peter Vosshall and Werner Vogels, ACM, 2007
- [5] <http://punetech.com/cloud-computing-and-high-availability/>
- [6] Bhavin Raichura, Ashutosh Agarwal, "Service Exchange@Cloud", Setlabs Briefings on Cloud Computing, Vol 7 NO 7, 2009
- [7] Bhavin Raichura, Vijay Rao, "Lean IT Transformation", ebizQ.net, 2009-03, <http://www.ebizq.net/topics/saas/features/11121.html>
- [8] Ian F. Akyildiz, Tommaso Melodia, Kaushik R. Chowdhury, "A survey on wireless multimedia sensor networks",
- [9] Nikos Dimokas, Dimitrios Katsaros, Yannis Manolopoulos, "Cooperative Caching in Wireless Multimedia Sensor Networks"
- [10] Adi Mallikarjuna Reddy V, Siva Prasad Katru, Kumar Padmanabh, "Can We Plug Wireless Sensor Network to Cloud?", Infosys Setlabs Briefings on Cloud Computing, Vol 7 NO 7, 2009
- [11] <http://www.oracle.com/technology/tech/cloud/index.html>
- [12] <http://contentoncontentmanagement.com/2008/02/29/amazon-s3-file-store-provider-for-oracle-ucm/>
- [13] Nam Doan-Huy, YiHong Xu, and Narshimha Rao Kondapaka, "Building Enterprise 2.0 Applications", Oracle Technology Network, Architect: Enterprise 2.0, August 2009
- [14] Demed L'Her, "SOA Suite on Amazon EC2", Oracle SOA Suite - Team Blog, October 15, 2009, http://blogs.oracle.com/SOA/2009/10/soa_suite_on_ec2_-_draft.html
- [15] "Oracle Innovations Behind Cloud Computing", This Week in Innovation, December 18, 2009, <http://www.oracle.com/innovation/feature-cloud-computing.html>
- [16] <http://www.oracle.com/technology/products/coherence/index.html>
- [17] "Platform-As-A-Service Private Cloud with Oracle Fusion Middleware", An Oracle White Paper, October 2009
- [18] "Architectural Strategies for Cloud Computing", An Oracle White paper in Enterprise Architecture, August 2009

About the Author

Nidhi Tiwari is a Senior Technology Architect in Software Engineering and Technology Labs, Infosys Technologies.

Nidhi has over ten years of industry experience in application design and development, performance engineering methodologies and cloud computing. She is currently involved in research and development activities of performance engineering and cloud computing streams. She interacts frequently with customers to understand their application performance challenges and advises on best practices for performance management. She has played a pivotal role in evolution and development of a performance modeling and simulation in-house tool. She has multiple publications and patents in performance engineering and cloud computing streams to her credit. She can be contacted at nidhi_tiwari@infosys.com

Did you know?

Infosys among the world's top 50 most respected companies

Reputation Institute's Global Reputation Pulse 2009 ranked Infosys among the world's top 50 most respected companies.



About Infosys

Many of the world's most successful organizations rely on Infosys to deliver measurable business value. Infosys provides business consulting, technology, engineering and outsourcing services to help clients in over 30 countries build tomorrow's enterprise.

For more information about Infosys (NASDAQ:INFY), visit www.infosys.com.

Global presence

Americas

[Brazil](#)
[Nova Lima](#)
[Canada](#)
[Calgary](#)
[Montreal](#)
[Toronto](#)
[Mexico](#)
[Monterrey](#)
[United States](#)
[Atlanta](#)
[Bellevue](#)
[Bentonville](#)
[Bridgewater](#)
[Charlotte](#)
[Detroit](#)
[Fremont](#)
[Hartford](#)
[Houston](#)
[Lake Forest](#)
[Lisle](#)
[New York](#)
[Phoenix](#)
[Plano](#)
[Quincy](#)
[Reston](#)

Asia Pacific

[Australia](#)
[Brisbane](#)
[Melbourne](#)
[Perth](#)
[Sydney](#)
[China](#)
[Shanghai](#)
[Hangzhou](#)
[Hong Kong](#)
[Central](#)
[India](#)
[Bangalore](#)
[Bhubaneshwar](#)
[Chandigarh](#)
[Chennai](#)
[Gurgaon](#)
[Hyderabad](#)
[Jaipur](#)
[Mangalore](#)
[Mumbai](#)
[Mysore](#)
[New Delhi](#)
[Pune](#)
[Thiruvananthapuram](#)
[Japan](#)
[Tokyo](#)
[New Zealand](#)
[Wellington](#)
[Philippines](#)
[Metro Manila](#)
[Singapore](#)
[Singapore](#)

Europe

[Belgium](#)
[Brussels](#)
[Czech Republic](#)
[Brno](#)
[Prague](#)
[Denmark](#)
[Copenhagen](#)
[Finland](#)
[Helsinki](#)
[France](#)
[Paris](#)
[Germany](#)
[Frankfurt](#)
[Stuttgart](#)
[Walldorf](#)
[Ireland](#)
[Dublin](#)
[Italy](#)
[Milano](#)
[Norway](#)
[Oslo](#)
[Poland](#)
[Lodz](#)
[Spain](#)
[Madrid](#)
[Sweden](#)
[Stockholm](#)
[Switzerland](#)
[Geneva](#)
[Zurich](#)
[The Netherlands](#)
[Amsterdam](#)
[United Kingdom \(UK\)](#)
[London](#)

Middle East and Africa

[Kingdom of Saudi Arabia](#)
[Riyadh](#)
[Mauritius](#)
[Reduit](#)
[UAE](#)
[Dubai](#)
[Sharjah](#)

For more information, contact askus@infosys.com

www.infosys.com

© 2011 Infosys Technologies Limited, Bangalore, India. Infosys believes the information in this publication is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of the trademarks and product names of other companies mentioned in this document.

Building 
Tomorrow's Enterprise