A Cloud Multimedia Platform

Dejan Kovachev and Ralf Klamma

Information Systems & Database Technologies, RWTH Aachen University
Ahornstr. 55, D-52056, Aachen, Germany
{kovachev|klamma}@dbis.rwth-aachen.de

Abstract. Social networking web applications such as Facebook and Flickr present new challenges for storing and processing user generated content, i.e. multimedia. Handling massive amounts of data requires special systems that need upfront investment, which may hinder the realization of new innovative ideas. Instead, cloud computing as a new emerging operations model promises to deliver elastic on-demand unlimited computing resources as a utility. In this position paper we propose architecture for a Cloud Multimedia Platform that does the heavy-lifting for massive amounts of multimedia storage and processing in the spirit of the cloud computing paradigm.

1 Introduction

Online social communities are an integral and important part of the Web. The Web has become a powerful and ubiquitous delivery channel for mass social interaction and collaboration applications. As a result, an enormous quantities of data are produced/consumed every day [1]. This data comes in forms of content (multimedia), structure (links) and usage (logs). Storing, managing, searching and delivering such data volumes introduce new challenges; thus motivating development of novel data systems.

Cloud computing envisions the notion of delivering software services and customizable hardware configurations to public access, similar how public utilities (electricity, water, etc.) are available to the common man [2]. The concept of cloud computing can include different computer technologies, such as networking infrastructure, Web 2.0, virtualization, SOA and other technologies. The cloud abstracts infrastructure complexities of servers, applications, data, and heterogeneous platforms, enabling users to plug-in at anytime from anywhere and utilize storage and computing services as needed at the moment.

Handling user generated content imposes problems more related to the data volumes, privacy, and delivery latency. Buying upfront the entire infrastructure storage and processing user generated content for a new social networking site can be problematic due to resource scarcity. Furthermore, setting up the computing environment can take additional resources, such as technical personal and time. However, it happens that sites do not go off. As an alternative, the cloud computing model fits naturally in such scenarios [3]. The cloud basically delivers the needed services to the users no matter the load, where charging happens

only by the used resources (CPU, storage, network bandwidth consumption), e.g. at Amazon's Elastic Compute Cloud. In the fast changing Web, delivery of new ideas is often crucial. The developers don't have the time, resources, or expertise to implement scalable infrastructure. Therefore, we believe that providing multimedia processing and storage in the cloud can leverage the dynamics of social networking sites, as there is an economic benefit of doing so.

2 A Cloud Computing Model Need

Consider an example of a social networking site. When the number of users on some social networking site starts to grow, then grow the amounts of user generated data.

An example for a multimedia processing leveraged in the cloud is a production of multiple versions of the same multimedia artifact. These versions could be different image sizes or image quality (thumbnails or mobile phone version). This task normally does not represent a difficult problem. However, in the case of social network site the size really matters (Facebook had 20 billion images as of 2009). That means processing the multimedia in reasonable time can be delivered by using the "unlimited" on-demand computing power of the cloud. Another example is image retrieval from large datasets. Finding the most similar image from a given set of known images requires feature extraction, such as color, texture, histogram, etc. Furthermore, it requires online real time feature comparison between the search item and the dataset.

We have assessed the needs of multimedia data management in cloud computing environment settings. As a result, we have extracted a set of principal requirements for the cloud multimedia platform:

R1: *Scalability.* Handling more data should be done by simply adding more machine instances; thus more power to the system.

R2: *Elasticity.* The infrastructure must be able to elastically adjust assigned resource (scale up as well scale down) according the usage rates.

R3: *Abstraction.* Developers are insulated from the details of provisioning servers, replicating data, recovering from failure, adding servers to support more load, securing data.

R4: *Simplicity.* There are simple interfaces for a development of new services and resource provisioning.

R5: *Interoperability.* The interoperability enables exporting from and importing in the cloud.

R6: *Distributed Data Management.* Large scale data is stored and managed in a distributed manner. This includes also processing and managing user generated content from a social network with millions of users spread around geographically.

**R7:** *Responsive Services.* The services respond timely and effectively regardless to the amount of data.

R8: *Service Orchestration.* Site administrators can orchestrate multiplex services to compose uniform flow to the user.

If we suppose that our social networking site was built on the cloud computing platform, and the site becomes popular, managing and operating the user generated content is handled with less problems.

3 Architecture for Cloud Multimedia Platform

We build a set of scalable, highly-available multimedia storage and processing services, that can be used across different applications. It is easy to build and rapidly evolve a Web-scale applications on top of the offered horizontal services. Developers focus on the applications logic, and the work around scaling and high availability is done in the cloud layer, instead in the application layer.

The architecture for the Cloud Multimedia Platform is shown on Fig.1. Since storage and computing on massive amounts of data are the key technologies for a cloud computing infrastructure [4], using virtual machines in combination with suitable middleware provide convenient solution for scalable computing systems. Our physical infrastructure is based on *Sun Solaris Containers* virtualization technology[1], which allows the user to allocate a system's various resources, such as memory, CPUs, and devices, into logical groupings and create multiple, discrete systems, each with their own operating system, resources, and identity within a single computer system. The virtualization layer based on the Solaris Containers technology enables provisioning of resources as needed. (Requirements R1, R2 and R3)
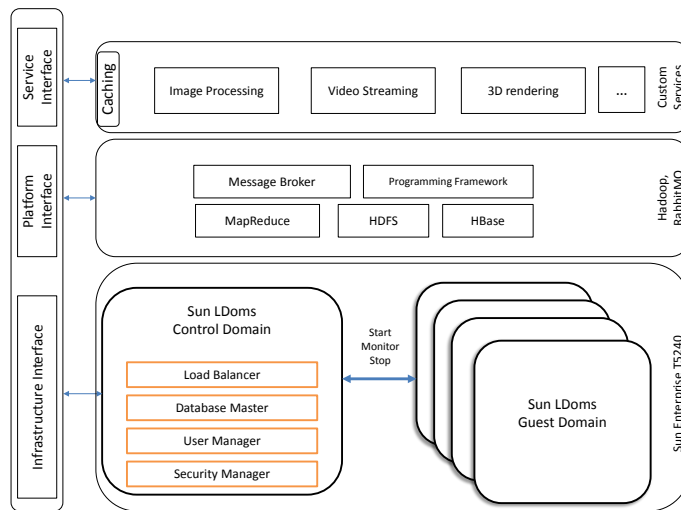
Fig. 1. Cloud computing architecture supporting multimedia services

---

[1] http://www.sun.com/software/solaris/virtualization.jsp

To ease the development of large scale data processing and storage applications we use the distributed computing framework *Hadoop*[2], which enables a cloud service similar to Platform as a Service (PaaS). The core of Hadoop are MapReduce, Hadoop Distributed File System (HDFS) and HBase. *MapReduce* [5] is one of the most popular programming paradigms for convenient large-scale computing on commodity hardware. *HBase* is a scalable, distributed database that supports structured data storage for large tables. (Requirements R3, R4 and R6)

Since we use many virtual machines running different disparate services, we use a messaging tier that helps tie the services work together. For example, we use *RabbitMQ*[3], which is an open source message broker based on the Advanced Message Queuing Protocol(AMQP).

For example, users produce a lot of multimedia content that is stored in HDFS. When some processing over the content is needed, such as feature extraction, transcoding, or resizing, Hadoop MapReduce can do it in a batch processing. Rankings, tags or other multimedia related metadata is stored within HBase.

4 Future Work

This work is only the first step of many. We plan to investigate which multimedia services are reasonable to be delivered through the cloud computing model. We need to expose simple interfaces to horizontal services, which are used among different vertical applications. The horizontal services in our cloud provide platforms to store, process and effectively deliver data to users. (Requirement R8) The interoperability can be enhanced by using standards for the multimedia metadata. (Requirement R5) We want to evaluate the effectiveness of the cloud computing model; we can easily test within our infrastructure the trade-off between putting more power on a single machine (scale-up) and putting multiple virtual machine instances (scale-out).

References

1. Baeza-Yates, R., Ramakrishnan, R.: Data challenges at Yahoo! In: EDBT '08, New York, NY, USA, ACM (2008) 652–655
2. Expert Group Report: The Future of Cloud Computing. Opportunities for European Cloud Computing Beyond 2010 (2010)
3. Armbrust, M., Fox, A., et al.: Above the Clouds: A Berkeley View of Cloud Computing. Technical report, EECS Department, University of California, Berkeley (February 2009)
4. Peng, B., Cui, B., Li, X.: Implementation Issues of A Cloud Computing Platform. IEEE Data Eng. Bull. **32**(1) (2009) 59–66
5. Dean, J., Ghemawat, S.: Mapreduce: simplified data processing on large clusters. In: OSDI'04, Berkeley, CA, USA, USENIX Association (2004) 10–10

---

[2] http://hadoop.apache.org
[3] http://www.rabbitmq.com