Review on Cloud Gaming Frameworks

Gajendra Sharma*

Department of Computer Science and Engineering, Kathmandu University, Nepal

Submission: January 31, 2019; **Published:** April 04, 2019

*Corresponding author: Gajendra Sharma, Department of Computer Science and Engineering, Kathmandu University, Public university in Dhulikhel, Nepal

Abstract

The cloud computing has evolved from the parallel and distributed computing, so that its use has been seen in wide range of scopes. Now days even gaming services are provided from cloud, from SaaS to GaaS (Gaming as a Service). This literature review overviews some general cloud gaming framework related to remote rendering, cognitive components-based gaming, GPU virtualization. Finally reach to a conclusion that distributed computing in modern GPU technologies have been boon to cloud computing, paving a way for the future of cloud gaming systems. Moreover, suggest that more research should be done keeping security of such cloud gaming framework in mind.

Keywords: Cloud computing; Gaming as a service; Virtualization; Low latency; Architectural framework; Cheating attack

Abbreviations: GaaS: Gaming as a Service; GA: Gaming Anywhere; VMs: Virtual Machines; QoE: Quality of Experience; GPU: Graphics Processing Unit

Introduction

The hardware limitation in our PC, laptop and mobile devices can be overcome by performing heavy computation intensive tasks remotely on more computationally powerful systems. Data centers is one of the options which can be used to offload computations off the mobile devices. Recently the use of cloud computing concept has increased the efficiency, scalability and elasticity of the datacenters. Utilization of widely deployed data centers by the cloud computing has provided new opportunities for new as well as existing applications. From file sharing and document synchronization to media streaming adopted by a variety of applications, notably web applications such as Google Docs and Dropbox, and video streaming by services like YouTube. Recently, advances in cloud technology have expanded to allow offloading not only of traditional computations but also of more complex tasks as 3D rendering thus making the Cloud Gaming a possibility. Modern Games have become highly computation intensive with complex graphics/AI logics/ Massive Multiplayer. The advent of cloud computing has changed the way of using internet for efficiency with more productivity. Offloading the computation intensive tasks to the cloud servers. The studies such as [1] have shown that with the help of cloud services a cloud gaming platform can reduce unnecessary actors from the business, thus depicting the evolution of a small software firm's product and business model, evolving from 15,000 users in 2005 to 3,000,000 users in 2010 with introduction of cloud computing concepts. The study also refers the cloud platform as a solution to piracy problems. Some have even found that the cloud platform gaming results in low energy consumption on user's end [2]. So, this paper overviews some of the studies on cloud gaming platform frameworks.

Literature Review

The online games of the past and the present such as standalone application games, browser based, uses the web servers for the retrieval and storage of user specific data in which all the core game engine resides in the client terminal. But the cloud computing has a brought about change in the framework from these concepts (Figure 1).

Figure 1: Basic components and steps in a game system.

Generally, a game can be considered as a software with the following components or modules: input, Game Logic/AI, Rendering and Display Output.

From most of the research [3-6] there are generally three general frameworks proposed for cloud gaming systems.

All game components on cloud

Among the three the two frameworks are shown in the Figure 2. In Figure 2 left framework, all the game logic

components along with the game instance are placed within the game servers in cloud. The client terminal is responsible to continuously take the user input from the input devices such as the controller, joystick, keyboard, mouse or touch screen, encode the input stream appropriately and send it to the game instance server in the cloud. The game logic processes the inputs and results in some change in the game environment,

the screen frame is rendered by the renderer which is actually job of the GPU (Graphics Processing Unit), the rendered screen is captured and encoded into video by encoder and sent to the video decoder unit in the client terminal which is streamed and played in the terminal. So, the client terminal generates the control signal and directs it to the cloud while cloud generates the visual and audio data to be sent to the terminal.
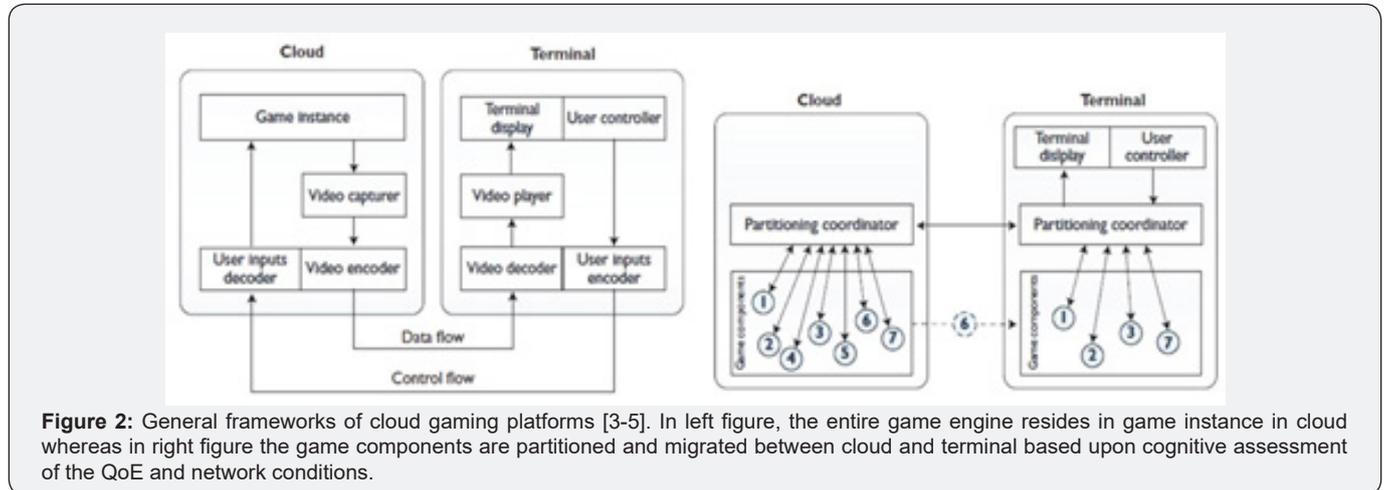

Figure 2: General frameworks of cloud gaming platforms [3-5]. In left figure, the entire game engine resides in game instance in cloud whereas in right figure the game components are partitioned and migrated between cloud and terminal based upon cognitive assessment of the QoE and network conditions.

There are certain advantages gained from this framework. The terminal device could be of a thin client form as well as platform independent, that is games could be played irrespective of the OS and hardware requirements of games, even on low powered devices such as Smartphones/Tablets providing a good platform for mobile cloud gaming. There is more efficiency for the terminal in terms of energy. Apparently, such architecture is used by almost all the cloud gaming services such as, on Live and Gai Kai (now owned by Sony in Playstation NOW), Nvidia Grid, G-Cluster etc. The Gaming Anywhere also uses similar framework [4], the first open cloud gaming system which performs multimedia streaming using open standard protocols (RTSP, RTP and RTCP). It was found that it provides 3 times lower processing delays and better video quality than OnLive gaming system. They also incorporated observer mode in which the server delivers encoded A/V frames using standard RTSP and RTP protocols so, an observer can observe a game play be accessing the corresponding game URL with multimedia players such as VLC player. Gaming Anywhere can be employed by the researchers, game developers, service providers, and end users for setting up cloud gaming testbeds.

Cognitive resource allocation

The second framework is an attempt to resolve the diversity in terminal devices and frequent changes in network quality of service and cloud responses which cause in the unstable Quality of Experience (QoE) for game players. So according to [6] the solution to the problem could be having a cognitive system to realize the situation of the resources and characteristics of cloud, the network and terminal devices so that the system can dynamically utilize the resources. In this the game are

modularized into abstract interdependent components which can migrate from the cloud to the user terminal which is called as on Loading process carried out by on Loading Manager.

The onloading process can either be performed before gaming session starts or be running in the background during the gaming session. And the components to be migrated are selected by the Partitioning Coordinator which is responsible for dynamic partitioning. These are all driven by continuously analyzing and evaluating the performance for maintaining an acceptable QoE for users. The components migration also aims at Partial Offline Execution of the game so that the player will not be disturbed by the temporary disconnection of Internet. Finally, the authors have shown that with the use of adaptive partitioning it is able to provide optimal solution in terms of overall latency in comparison to the case in which all the components are either reside in cloud and in the client terminal. From the paper, it is evident that the most important element of the cognitive framework is the onloading and partitioning process. So, there are a lot of future research in these elements may improve the efficiency and specially the security of the system.

Virtualization of GPUs

In these both frameworks, it is evident that when using the cloud services, each game instance will run in a Virtual Machine (VM) generated by virtualization. Consolidating multiple virtual machines (VMs) onto a physical machine, so as to virtualize various resources, including CPUs, networks, storages, and GPUs for sharing. While virtualizing CPUs, network interfaces, and storages is rather mature, virtualizing GPUs is still considered experimental. Such consolidation

should be carefully done since it can degrade the performance. Several papers such as [7,8] have demonstrated degraded performances in terms of lower frame rate, high response time and low video quality when virtualizing GPU and shared among multiple VMs. Keeping this in mind several other papers have tried newer technologies such as modern GPUs [9] such as Nvidia K2 and GRID based cloud system [9,10] is directed towards answering whether the modern GPUs are ready for cloud gaming. They found that the modern GPUs have low consolidation overhead and are powerful enough to concurrently support multiple GPU-intensive cloud games, the cloud gaming server with the modern GPUs may become CPU-bounded. Hence, offloading the software video codec to the GPUs can further improve the gaming experience. They also found that the mediated pass-through GPU virtualization implemented in the latest GPUs enables efficient GPU sharing among multiple VMs resulting in better performance than dedicated GPUs as well as the shared GPUs are scalable to a number of VMs (Figure 3).
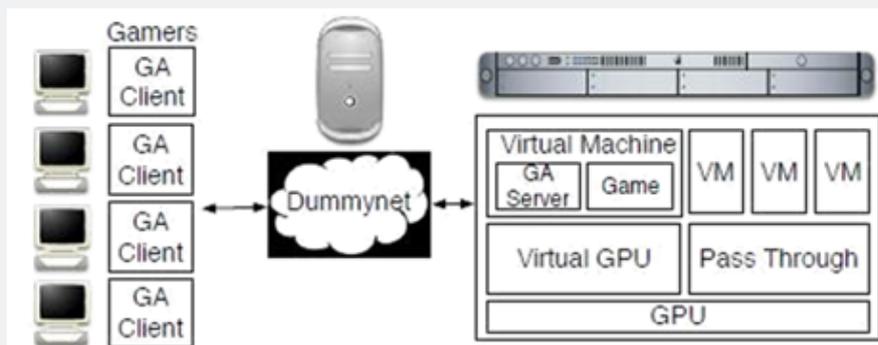

Figure 3: Architecture from [9] consisting of a cloud gaming server running multiple GA servers and games, a dummynet router, and several GA clients.

The other paper [10] is centered towards use of a concurrent server built up based on the integration of Nvidia Grid GPUs and Gaming Anywhere open source cloud gaming platform in order to reduce the delay in limited bandwidth and improve multi-client concurrent access efficiency.

NVIDIA GRID [11] is a series of products which includes GPU virtualization, remote processing and session management libraries, it allows multiple users to take advantage of GPU while experiencing graphics-intensive desktop applications and games. The authors have modified the game video streaming capture mode of the Gaming Anywhere (GA) and added support for concurrent accesses to achieve the largest number of games running in one gaming server. The authors observed that for the original GA it takes about 61ms to capture the game screen, convert the data format and compress the data, all stages processed by CPU while using the NVIDIA GPU, data format conversion and data compression H.264 is completed in about 4ms. This result was due to hardware acceleration of image format conversion and H264 compression in GPU (Figure 4).
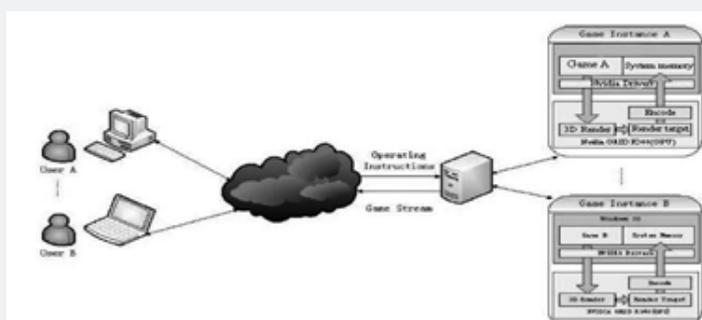

Figure 4: The architecture diagram of cloud gaming framework using virtualization of Nvidia GRID GPUs [10].

There are also a number of possible improvements made in the future which includes optimizing the capture mechanism to support more type of games, the concurrency to support more client connection and the data transfer to make the transfer process more stable. Recently, Nvidia too have released their own cloud gaming (streaming) service called Nvidia Grid for their gaming console Nvidia Shield [12].

Conclusion

Apart from the present online game architecture there are two frameworks for cloud gaming platform, one involving processing and rendering the entire game in cloud and then playing the captured video in client terminal, and the other involving cognitive concept of partitioning the game components and migrating them between the client terminal

and the cloud for maintaining acceptable QoE. We even came across some architecture with virtual machine not only with virtual CPU, storage, network only but with virtual GPU as well and how modern GPU technologies are helping to reduce the delay latencies and serve numerous concurrent users. So, the cloud gaming has a good future with the advancement in GPU technologies, specially related to its virtualization and concurrent access. GPUs have created a great platform for cloud gaming.

Also, security is a great concern for any distributed systems. This study [13] shows some security issues in any online games as listed in Table 1, presents common cheating methods in online games, dividing them into those which are specific to or of special relevance to online games and which are general to all network applications. But there are no studies made directed towards the security of the system as well as the users. For example, the cognitive framework keeps some game components in cloud as well as some in the terminal for dynamically adapting to the changing network and user experience conditions as well as partial offline gameplay.

But it should be kept in mind that malicious users could easily manipulate the local components than the remote ones by accessing the system memory using tools such as Cheat Engine [14] or manually using low level programming routines. Even in the other framework where entire game resides in cloud and rendered screen are streamed to the terminal, the users could still send input signals to the cloud game instance which indicates some potential of cheating and hacking through the input streams. There are also chances of packet injection and code injection techniques to successfully carry out security attacks regarding cheating and hacking the gaming system.

There are some research opportunities regarding how to make the cloud gaming systems resistant to these security issues related with cloud gaming frameworks. At least, some attack or cheat detection mechanism researches could be employed. Attempts should be made to maintain acceptable QoE for users even with such attack and cheat detection mechanism in the frameworks because adding the detection or prevention mechanism may introduce additional processing delays hence affecting the overall QoE.

References

1. Ojala A, Tyrvainen P (2011) Developing Cloud Business Models: A Case Study on Cloud Gaming. IEEE Software 28(4): 42-47.

2. Chun-Ying Huang, Yu-Ling Huang, Yu-Hsuan Chi, Kuan-Ta Chen, Cheng-Hsin Hsu (2015) To Cloud or Not to Cloud: Measuring the Performance of Mobile Gaming. In: Proceedings of the 2nd Workshop on Mobile Gaming (Mobi Games '15), ACM, New York, USA.

3. Wei Cai, Min Chen, Leung VCM (2014) Toward Gaming as a Service. IEEE 18(3): 12-18.

4. Chun-Ying Huang, Cheng-Hsin Hsu, Yu-Chun Chang, Kuan-Ta Chen (2013) Gaming Anywhere: an open cloud gaming system. In: Proceedings of the 4th ACM Multimedia Systems Conference (MMSys '13), ACM, New York, USA, p. 36-47.

5. K Lee, D Chu, E Cuervo, Y Degtyarev, S Grizan (2015) Out a time: Using speculation to enable low-latency continuous interaction for mobile cloud gaming. Mobi Sys 13th Internation Conference on Mobile Systems, Applications and Services.

6. Wei Cai, Conghui Zhou, Victor CM Leung, Min Chen (2013) A Cognitive Platform for Mobile Cloud Gaming. 5th International Conference on Cloud Computing Technology and Science, IEEE p. 72-79.

7. M Dowty, J Sugerman (2009) GPU virtualization on VMware's hosted I/O architecture. ACM SIGOPS Operating Systems Review 43(3): 73-82.

8. R Shea, J Liu (2013) On GPU pass-through performance for cloud gaming: Experiments and analysis. In: Proc. of ACM Annual Workshop on Network and Systems Support for Games (NetGames'13), Denver, CO.

9. Hua-Jun Hong, Tao-Ya Fan-Chiang, Che-Run Lee, Kuan-Ta Chen, Chun-Ying Huang (2014) GPU consolidation for cloud games: Are we there yet? Network and Systems Support for Games (NetGames), 13th Annual Workshop, p. 1-6.

10. Qingdong Hou, Chu Qiu, Kaihui Mu, Quan Qi, Yongquan Lu (2014) A Cloud Gaming System Based on NVIDIA GRID GPU. Distributed Computing and Applications to Business, Engineering and Science (DCABES), 13th International Symposium, p.73.

11. http://www.nvidia.com/object/cloud-gaming.html

12. http://shield.nvidia.com/

13. Yan J, Randell B (2009) An Investigation of Cheating in Online Games. IEEE 7(3): 37-44.

14. www.cheatengine.org