Cloud VR

Secure, Fast and Distributed Virtual Reality Solutions

Leon Koster

A thesis presented for the degree of
Bachelor of Sciences

Academie Creative Technology (ACT)
Saxion University of Applied Sciences
Netherlands
May 2020

1 Acknowledgements

I would like to acknowledge my university and teachers, Matthijs van Veen, Yiwei Jiang and Hester van der Ent, for their help with the creation of this paper.

2 Abstract

Contents

3 List of Figures and Tables

List of Figures

List of Tables

4 List of Abbreviations

Acronyms

FoV Field-of-View. 7, 10

HMD Head Mounted Display. 7, 8, 9, 10

ms Milliseconds. 5, 8, 9, 10, 12

MTP Motion-to-Photon. 9, 10, 12

RGB Red Green Blue. 5

UX User Experience. 9

VR Virtual Reality. 3, 4, 6, 7, 8, 9, 10, 12

5 Introduction

Recent developments in the field of Virtual Reality (VR) offer all kinds of opportunities in the field of training and entertainment. For training purposes, the audiovisual entry into a virtual world is where the biggest value is. The capabilities of artificial environments allow users to manage scenarios and experiences that cannot be simulated in the real world. VR also allows users to access the virtual training at any time and less physical facilities are required for exercises. Examples VR experiences include training maintenance at high altitudes (such as windmills), working under heavy loads and weather conditions in construction (Strukton) or maintenance on naval ships (Thales). These companies (and more) form the Industrial Reality Hub, which is one of the stakeholders of this project.

6 Preliminary Problem Statement

One of the essentials for a good Virtual Reality (VR) experience is a powerful computer system to render semi-realistic worlds. However, there are two problems here. First, this type of system is not available in every location. Certainly if realistic images have to be rendered in the simulation, it requires specialized and expensive machines that are difficult to move.

The second problem is that for rendering the VR training scenario, all kinds of data about the scenarios need to be available on the system. This can pose a problem when it concerns sensitive information, for example about all kinds of information defense systems or business sensitive information.

Focusing on these problems will lay the foundation for future research, to make CloudVR streaming a mature technology.

The aim of this project is to investigate the feasability of a streaming based VR approach with current cutting edge technology. Qualitative research methods will be used to gain in-depth insights about existing solutions and the current state of research into this topic. The data will be contextualized via a literature review of recent research papers and an analysis of the capabilites of existing solutions when applied to the research problem.

7 Problem Analysis

Together with the companies from the Industrial Reality Hub mentioned in the Introduction, Saxion wants to investigate how virtual reality can be rendered in the cloud in a safe and efficient manner. This involves looking at state-of-the art technology in the field of virtual reality, cloud computing, rendering and machine learning for one complete CloudVR pipeline. There are four research objectives here:

7.1 Architecture for a cloud VR system

One of the questions to be answered is what the CloudVR architecture should look like in terms of hardware and software. This not only concerns the servers, but also whether there are local ones rendering is required (see the next point).

7.2 Latency

Current market players such as Google Stadia (Google, 2019), GeForce Now (Nvidia, 2020c) and Xbox xCloud (XBox, 2019) already offer cloud gaming services that stream games over the internet. Powerful servers are used for rendering games that are then streamed to users in real time. A bottleneck with this technology is the latency (delay). This is because user input is first sent to a server, which renders these new images, after which they are sent back to the users, all without disturbing them. The mentioned platforms all use network optimization. Low latency is very important for VR, where head movements should be converted to images in under 20 Milliseconds (ms), to prevent motion sickness (Abrash, 2012). The research for techniques for reducing latency is one of the spearheads of the CloudVR project. The following research directions are relevant here:

Network optimization As with the platforms described above, network optimization is one of the techniques which needs to be investigated. The question is to what extent an optimized network can reduce latency and how it relates to the quality of the network connection.

Two-step rendering One of the options to bypass latency is to render in two steps. The delay is not so much reduced, but avoided. The server renders next to RGB also positions and BRDF variables for each pixel. Afterwards on the user's (less powerful) hardware adjustments are made so that the image corresponds to the current position of the user. By sending additional data, the user's local client can extrapolate the correct information and construct a frame that represents the correct head position in the last frame, meanwhile it is waiting for the correct next frame from the server.

Behavioral prediction Another possibility to reduce latency is by predicting user input through machine learning. This will mainly revolve around it analyzing head movements to find out what behavior can be expected. With this information we can render any part of the virtual world before it is viewed by users. If this information is then forwarded from the cloud to the location of the VR experience, what information is displayed can be selected on the spot.

7.3 Multi-user experiences

One of the questions with a CloudVR solution is how to deal with multi-user VR experience where users at another location share a VR experience via a network. The interaction with each other and the environment are a point of attention.

7.4 GPU scaling

One of the advantages of cloud rendering is that in theory it gives the possibility of unlimited computing capacity. This gives the opportunity to all kinds of touristic feats (graphic), and interaction (physics). It is therefore interesting as part of the CloudVR pipeline to investigate how techniques such as NVLink and NVSwitch 5 (Nvidia, 2020b) could be used for high-quality VR experiences.

8 Theoretical Framework

In order to thoroughly understand the aim and subject of the research, it is important to explore different existing solutions and literature. Therefore, the subjects that will be discussed in the following theoretical framework are Cloud Streaming/Cloud Computing and Virtual Reality. Within this theoretical framework, definitions of the subjects will be given as well as current insights into these subjects. The topics reflect knowledge needed to understand the problem space. Together all of the topics make up the 360 scan.

8.1 Cloud Streaming/Cloud Computing

8.1.1 Definition

According to Armbrust et al. (2010) Cloud computing is defined as follows: "Cloud computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the data centers that provide those services." (Armbrust et al., 2010) We can then further define Cloud streaming as the applications that are delivered over the internet as a service.

8.1.2 Existing Solutions and Technology

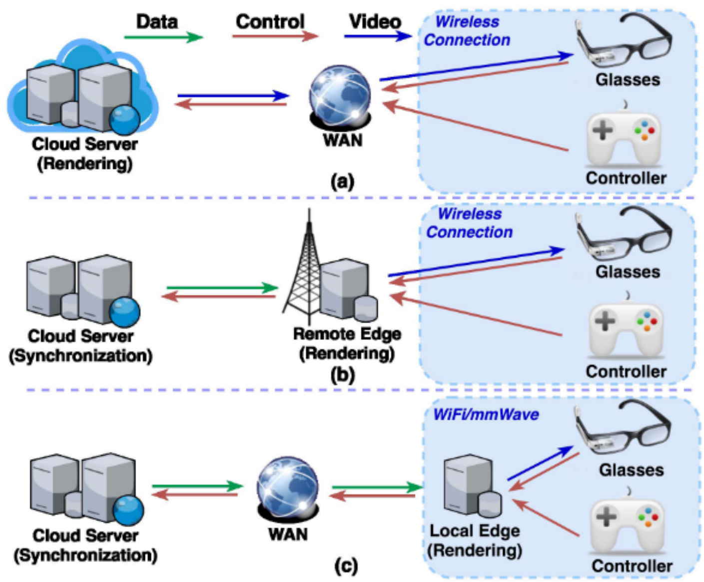
Several commercial gaming Cloud streaming services already exist, such as Google Stadia (Google, 2019), XBox XCloud (XBox, 2019) and Nvidia GeForceNow (Nvidia, 2020c). These applications deliver conventional games from a powerful computer in a server to the client device at home. Despite initial setbacks, cloud streaming is now a mainstream technology. The start of 2020 also saw the first experimental cloud VR streaming development kits, such as Nvidia's CloudXR (Nvidia, 2020a), and closed beta's for commercial cloud VR streaming services (Shadow, 2020). There is also a variety of Infrastructure-as-a-service (IaaS) platforms, such as Amazon's AWS (Amazon, n.d.), Microsoft's Azure (Microsoft, n.d.) and Google's Cloud Platform (Google, n.d.-a), that provide generic computing power and storage in a cloud computing/streaming context. For more information about these applications and technologies, please refer to the Literature Analysis.

8.1.3 System Architecture (for a cloud VR system)

One of the main considerations when designing a cloud VR streaming application is the decision to either use a Cloud, Remote Edge or Local Edge computing device for the rendering of the frames (See Figure 1 and Hou et al., 2017):

- A cloud server renders the Field-of-View (FoV) (current view) remotely and streams the corresponding video to the user's Head Mounted Display (HMD).

Figure 1: Cloud Server, Remote Edge, Local Edge visualized (Hou et al., 2017)



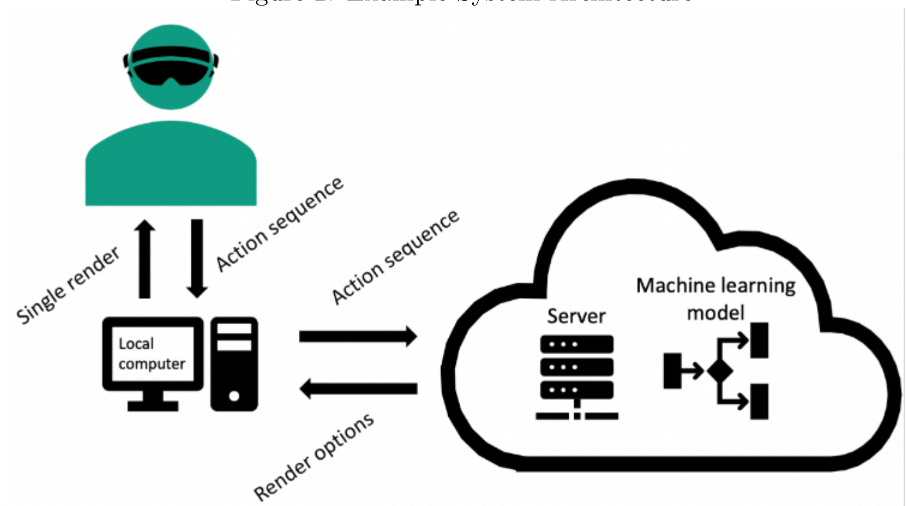
- A Remote Edge sever recieves multiple views that are rendered remotely on cloud servers, stitches them together to a 360-degree video, and streams the video to the user's HMD
- A Local Edge server recieves compressed models as well as textures, renders it locally and streams the video to the user's HMD.

For the purposes of this research I will ignore Local Edge system architectures. The reason its that one of the reasons for this research paper was the desire to keep data as safe as possible, which in this case means keeping in the cloud. A Local Edge system by design requests and recieves business data to render the frame for the user locally. For this reason a Local Edge approach would be the wrong direction to research in. An example architecture of a cloud VR solution that keeps the business data in the cloud can be seen in Figure 2.

8.1.4 Latency

The most important metric for a system architecture is the latency between the user input, such as movement of the HMD, and the updated frame appearing on the users display. Recent measurements of cloud gaming services measure this latency at between 135 and 240 ms (Chen et al., 2019). This is acceptable for most games, except maybe high intensity reaction games. VR unfortunately has severely stricter latency requirements, which are elaborated upon in the next section.

Figure 2: Example System Architecture



8.2 Constraints of Virtual Reality

As mentioned before, when developing a VR application, there are a few physical constraints that developers need to be aware of. The most important threshold to know is the 20 ms Motion-to-Photon (MTP) delay. Upon input from the Head Mounted Display (HMD), the developer has to display a new rendered image within an average of 20 ms to avoid motion sickness for users. The more this threshold can be undercut, the better the chances to have an acceptable gameplay experience without motion sickness. Interaction input, such as the input from the controllers, can safely be processed at delays of >100 ms without any negative repercussions in terms of User Experience (UX). For more information, please refer to the Literature Analysis.

9 Literature Review

9.1 Modern (<5 years old) research and technology

Within the last decade the cloud computing space has expanded rapidly and with it the possibilities. Today, even individuals can set up an experimental cloud Virtual Reality (VR) streaming solution from pre-made components (TayoEXE, 2019), such as the service from cloud computing company Shadow (Shadow, 2015), who recently also announced a closed beta for their dedicated VR streaming services (Shadow, 2020). The most recent years have also seen the rise of commercial cloud streaming for gaming services, such as Google Stadia (Google, 2019), XBox XCloud (XBox, 2019) and Nvidia GeForceNow (Nvidia, 2020c), but not without problems, as the services were quickly overwhelmed on launch. Modern video compression codecs, like the AV1 codec introduced in 2018 (AllianceForOpenMedia, n.d.), are getting better and better at compressing high-resolution video streams and together with an application like WebRTC (Google, n.d.-b) which offers latency optimizations via peer-to-peer networking and more, they lay the foundation for modern cloud streaming applications. Technologies like Google's Seurat Image-Based Scene Simplification System (Google, 2018) offer even further optimizations in areas other than networking and transmitting data. With ever increasing performance and optimizations, the feasibility of mainstream cloud VR streaming is just a matter of time.

Research Papers like the ones from Liu et al., 2018 or from Shi et al., 2018 demonstrate the viability and technical feasibility of cloud VR streaming . They developed innovative solutions to achieve and undercut the 20 Milliseconds (ms) Motion-to-Photon (MTP) delay barrier while streaming VR content. 20 ms is the agreed upon threshold a frame should have, from receiving the input to displaying the frame on the Head Mounted Display (HMD), to avoid inducing motion sickness (Abrash, 2012). One such solution is a low latency control loop that streams VR scenes containing only the user's Field-of-View (FoV) and a latency adaptive margin area around the FoV. This allows the clients to render locally at a high refresh rate to accommodate and compensate for the head movements before the next motion update arrives. (Shi et al., 2018). Another angle of attack leverage's the power of parallel rendering, encoding, transmission and decoding, together with a Remote VSync Driven Rendering approach to minimize MTP latency (Liu et al., 2018). The prototype for this experiment was based on commodity hardware, which further demonstrates the feasibility of cloud VR streaming.

10 Final Problem Statement

11 Research Questions

Main Question 1:

What is the current state of cloud Virtual Reality (VR) streaming?

Sub Question 1: How (in)effective are existing solutions when applied to the cloud streaming VR context?

Sub Question 2: What research has been done on the shortcomings from question 1 ? Which area still requires the most research ?

Sub Question 3: What are the most important considerations when designing an architecture for cloud VR streaming?

Sub Question 4: What methods are the most efficient way to reduce Motion-to-Photon (MTP) Latency to $\leq$20ms?

12 Methodology

13 Experiments

14 Results

15 Discussion

16 Conclusion

References

Bryson, S. (1996). Virtual reality in scientific vizualisation. *Communications of the ACM*, *39*(5).

Burdea, G. C., & Coiffet, P. (2003). *Virtual reality technology.*

Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., & Zaharia, M. (2010). A view on cloud computing. *Communications of the ACM*, *53*(4), 50–58. https://doi.org/https://dl.acm.org/doi/fullHtml/10.1145/1721654.1721672

Abrash, M. (2012). *Latency – the sine qua non of ar and vr.* http://blogs.valvesoftware.com/abrash/latency-the-sine-qua-non-of-ar-and-vr/ accessed: 14.05.2020

Shadow. (2015). *Transform your device into an gaming pc.* https://shadow.tech/usen accessed: 08.05.2020

Hou, X., Lu, Y., & Dey, S. (2017). Wireless ar/vr with edge/cloud computing. *26th International Conference on Computer Communication and Networks (ICCCN).* https://doi.org/10.1109/ICCCN.2017.8038375

Google. (2018). *Seurat: A scene simplification technology designed to process very complex 3d scenes into a representation that renders efficiently on mobile 6dof vr systems.* https://github.com/googlevr/seurat accessed: 14.05.2020

Liu, L., Zhong, R., Zhang, W., Liu, Y., Zhang, J., Zhang, L., & Gruteser, M. (2018). Cutting the cord: Designing a high-quality untethered vrsystem with low latency remote rendering. *Proceedings of MobiSys ACM '18.* https://doi.org/https://doi.org/10.1145/3210240.3210313

Shi, S., Gupta, V., Hwang, M., & Jana, R. (2018). Mobile vr on edge cloud: A latency-driven design. *MMSys '19.* https://doi.org/https://doi.org/10.1145/3304109.3306217

Chen, S.-W. (.-T., Chang, Y.-C., Tseng, P.-H., Huang, C.-Y., & Lei, C.-L. (2019). Cloud gaming latency analysis: Onlive and streammygame delay measurement. https://www.iis.sinica.edu.tw/~swc/onlive/onlive.html

Google. (2019). *Stadia - one place for all the ways we play.* https://stadia.google.com/ accessed: 11.05.2020

TayoEXE. (2019). *Test: No vr-ready pc required, cloud pc vr streaming via shadow and virtual desktop quest.* https://www.reddit.com/r/OculusQuest/comments/c5xnux/test_no_vrready_pc_required_cloud_pc_vr_streaming/ accessed: 08.05.2020

XBox. (2019). *Project xcloud.* https://www.xbox.com/en-US/xbox-game-streaming/project-xcloud accessed: 11.05.2020

Nvidia. (2020a). *Nvidia cloudxr sdk.* https://developer.nvidia.com/nvidia-cloudxr accessed: 13.05.2020

Nvidia. (2020b). *Nvlink & nvswitch: Advanced multi-gpu systems.* https://www.nvidia.com/en-us/data-center/nvlink/ accessed: 14.05.2020

Nvidia. (2020c). *Your games. your devices. play anywhere.* https://www.nvidia.com/en-us/geforce-now/ accessed: 11.05.2020

Shadow. (2020). *Vr explorer log: Shadow's entrance into vr begins!* https://community.shadow.tech/usen/blog/news/vr-explorer-log accessed: 14.05.2020

AllianceForOpenMedia. (n.d.). *Av1 features*. https://aomedia.org/av1-features/ accessed: 12.05.2020

Amazon. (n.d.). *Amayon web services (aws) - cloud computing*. https://aws.amazon.com/ accessed: 12.05.2020

Google. (n.d.-a). *Google cloud computing services*. https://cloud.google.com/ accessed: 12.05.2020

Google. (n.d.-b). *Webrtc*. https://webrtc.org/ accessed: 12.05.2020

Microsoft. (n.d.). *Cloud computing services — microsoft azure*. https://azure.microsoft.com/en-us/ accessed: 12.05.2020

17 Appendices