

ICNRG Working Group
Internet-Draft
Intended status: Informational
Expires: January 15, 2019

C. Westphal
Huawei
July 14, 2018

AR/VR and ICN
draft-westphal-icnrg-arvr-icn-00

Abstract

This document describes the challenges of AR/VR in ICN.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 15, 2019.

Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Internet-Draft

ICN-ARVR

July 2018

Table of Contents

1.	Introduction	2
2.	Definitions	3
2.1.	Use Cases	4
2.1.1.	Office productivity, personal movie theater	4
2.1.2.	Retail, Museum, Real Estate, Education	4
2.1.3.	Sports	4
2.1.4.	Gaming	5
2.1.5.	Maintenance, Medical, Therapeutic	5
2.1.6.	Augmented maps and directions, facial recognition, teleportation	5
3.	Information-Centric Network Architecture	6
3.1.	Native Multicast Support	6
3.2.	Caching	7
3.3.	Naming	7
3.4.	Privacy	7
3.5.	Other benefits?	7
3.6.	Security Considerations	7
4.	References	8
4.1.	Normative References	8
4.2.	Informative References	8
	Author's Address	8

[1.](#) Introduction

Augmented Reality and Virtual Reality are becoming common place. Facebook and YouTube have deployed support for some immersive videos, including 360 videos. Many companies, including the aforementioned Facebook, Google, but also Microsoft and others, are offering devices to view virtual reality, ranging from simple mechanical additions to a smart phone, such as Google Cardboard to full fledged dedicated devices, such as the Oculus Rift.

Current networks however, are still struggling to deliver high quality video streams. 5G Networks will have to address the challenges introduced by the new applications delivering augmented reality and virtual reality services. However, it is unclear that without architectural support, it will be possible to deploy such applications.

Most surveys of augmented reality systems (say, [[van2010survey](#)]) ignore the potential underlying network issues. We attempt to

present some of these issues in this paper. We also intend to explain how an Information-Centric Network architecture is beneficial for AR/VR. Information-Centric Networking has been considered for enhancing content delivery by adding features that are lacking in an

IP network, such as caching, or the requesting and routing of content at the network layer by its name rather than a host's address.

[2.](#) Definitions

We provide definitions of virtual and augmented reality (see for instance [[van2010survey](#)]):

Augmented Reality: an AR system inserts a virtual layer over the user's perception of the real objects, which combines both real and virtual objects in such a way that they function in relation to each other, with synchronicity and the proper depth of perception in three dimensions.

Virtual Reality: a VR system places the user in a synthetic, virtual environment with a coherent set of rules and interactions with this environment and the other participants in this environment.

Virtual reality is immersive and potentially isolating from the real world, while augmented reality inserts extra information onto the real world.

For the purpose of this article, we restrict ourselves to the audio-visual perception of the environment (even though haptic systems may be used) as a first step. Many of the applications of augmented and virtual reality similarly start with eyesight and sounds only.

Most of the AR/VR we consider here focuses on head-mounted displays, such as Oculus Rift or Google Cardboard.

There are obvious observations derived from these descriptions of virtual and augmented reality. One is that virtual reality only really needs a consistent set of rules for the user to be immersed into it. It could theoretically work on a different time scale, say where the reaction to motion is slower than in the real world. Further, VR only needs to be self-consistent, and does not require

synchronization with the real world.

As such, there are several levels of complexity along a reality-virtuality continuum. For the purpose of the networking infrastructure, we will roughly label them as 360/immersive video, where user is streaming a video stream with a specific viewing angle and direction; virtual reality environment, where the user is immersed in a virtual world and has agency (say, decide of the direction of the motion, in addition to deciding of the direction of her viewing angle); and augmented reality where the users' view is overlaid on top of the actual real view of the user.

The last application requires identifying the environment, generating and fetching the virtual artifacts, layering these on top of the reality in the vision of the user, in real time and in synchronization with the space dimensionality and the perception of the user, and with the motion of the user's field of vision. Such processing is very computationally heavy and would require a dedicated infrastructure to be placed within the network provider's domain.

[2.1.](#) Use Cases

For AR/VR specifically, there is a range of scenarios with specific requirements. We denote a few below, but make no claim of exhaustivity: there are plenty of other applications.

[2.1.1.](#) Office productivity, personal movie theater

This is a very simple, canonical use case, where the headmounted device is only a display for the workstation of the user. This has little networking requirements, as all is collocated and could even be wired. For this reason, it is one of the low hanging fruits in this space. The main issue is of display quality, as the user spends long hour looking at a screen, with a resolution, a depth of perception, and a reactivity of the headmounted display that should be comfortable for the user.

[2.1.2.](#) Retail, Museum, Real Estate, Education

The application recreates the experience of being in a specific area,

such as a home for sale, a classroom or a specific room in a museum. This is an application where the files may be stored locally, as the point is to replicate an existing point of reference, and this can be processed ahead of time.

Issues become then how to move the virtual environment onto the display. Can it be prefetched ahead of time; can it be distributed and cached locally near the device; can it be rendered in the device?

[2.1.3.](#) Sports

This attempts to put the user in the middle of a different real environment, as in the previous case, but adds to it several dimensions: that of real time, as the experience must be synchronized with a live event; that of scale, as many users may be attempting to participate in the experience simultaneously.

These new dimensions add some corresponding requirements, namely how to distribute live content in a timely manner that still corresponds

Westphal

Expires January 15, 2019

[Page 4]

Internet-Draft

ICN-ARVR

July 2018

to the potentially unique viewpoint of each of the users; how to scale this distribution to a large number of concurrent experiences. The viewpoint in this context also may impose different requirements, if it is that of a player in a basketball game, or that of a spectator in the arena. For instance, in the former case, the position of the viewpoint is well defined by that of the player, while in the latter, it may wildly vary.

[2.1.4.](#) Gaming

Many games place the user into a virtual environment, from Minecraft to multi-user shooter game. Platform such as Unity 3D allow creation of virtual worlds. Unlike the previous use case, there are now interactions in between the different participants in the virtual environment. This require communication of these interactions in between peers, and not just from a server onto the device. There are issues of consistency across users and synchronization issues.

[2.1.5.](#) Maintenance, Medical, Therapeutic

There exist a few commercial products where the AR is used to overlay instructions on top of some equipment so as to assist the agent in

performing maintenance. Surgical assistance may fall in this category as well.

The advantage of a specific task is that it facilitates the pattern recognition and the back-end processing as it is narrowed down. However, the requirements to overlay the augmented layer on top of the existing reality puts stringent synchronization and round-trip time requirements, both on the display and on the sensors capturing the motion and position.

2.1.6. Augmented maps and directions, facial recognition, teleportation

The more general scenario of augmented reality does not focus on a specific, well defined application, but absorbs the environment as observed by the user (or the user's car or the pilot's plane, if the display is overlayed on a windshield) and annotates this environment, for instance to specify directions. This includes recognizing patterns and potentially people with the help of little context beyond the position of the user. Another main target of AR is telepresence, where a person in a remote location could be made present, as if in another location, say with others in the same conference room. Teleportation plus the display of the workstation of a user (as in the first scenario above) may allow remote collaboration on enterprise tasks.

3. Information-Centric Network Architecture

We now turn our attention to the potential benefits that Information-Centric Networks can bring to the realization of AR/VR.

The abstractions offered by an ICN architecture are promising for video delivery. [RFC7933](#) [[RFC7933](#)] for instance highlights the challenges and potential of ICN for adaptive rate streaming. As VR in particular may encompass a video component, it is natural to consider ICN for AR/VR.

There is a lot of existing work on ICN (say, caching or traffic engineering [[su2013benefit](#)]) which could be applied to satisfy the QoS requirements of the AR/VR applications, when possible.

[3.1.](#) Native Multicast Support

One of the key benefits from ICN is the native support for multicast. For instance, [[macedonia1995exploiting](#)] quotes: "if the systems are to be geographically dispersed, then highspeed, multicast communication is required." Similarly, [[frecon1998dive](#)] states that: "Scalability is achieved by making extensive use of multicast techniques and by partitioning the virtual universe into smaller regions."

In the sport use case, many users will be participating in the same scene. They will have potentially distinct point of views, as each may look into one specific direction. However, each of these views may share some overlap with the others, as there is a natural focus point within the event (say, the ball in a basketball game).

This means that many of the users will request some common data and native multicast significantly reduces the bandwidth and in the case of ICN, without extra signaling.

Further, the multicast tree should be adhoc, and dynamic to efficiently support AR/VR. Back in 1995, [[funkhouser1995ring](#)] attempted to identify the visual interactions in between entities representing users in a VE so as to "reduce the number of messages required to maintain consistent state among many workstations distributed across a wide-area network. When an entity changes state, update messages are sent only to workstations with entities that can potentially perceive the change i.e., ones to which the update is visible.}" [[funkhouser1995ring](#)] was able to reduce the number of messages processed by client workstations by a factor of 40.

It is unclear that ICN can assist in identifying which workstations (or nowadays, which users) may perceive the status update of another user (but naming the data at the network layer may help). Nonetheless, the multicast tree to reach the set of clients that would require an update is dynamically modified and the support for multicast in ICN definitely supports this dynamic behavior.

[3.2.](#) Caching

The caching feature of ICN allows prefetching of data near the edge some of the more static use cases; further, in the case of multiple users sharing a VE, the caching allows to perform the content placement phase for some users at the same time as the content distribution phase of others, thereby reducing bandwidth consumption.

Caching is a prominent feature in an AR system: the data must be nearby to reduce the round-trip time to access the data. Further, AR data has a strong local component and therefore caching allows to keep the information within the domain where it will be accessed.

ICN naturally supports caching, and provides content-based security to allow any edge cache to hold and deliver the data.

[3.3.](#) Naming

Since only a partial Field of View is accessed from the whole spherical view at any point in time, tiling the spherical view into smaller areas and requesting the tiles that are viewed would reduce the bandwidth consumption of AR/VR systems. This raises the obvious question of naming semantics for tiles. New naming schemes that allow for tiling should be devised.

[3.4.](#) Privacy

By enabling caching at the edge, ICN enhances the privacy of the users. The user may access data locally, and thereby will not reveal information beyond the network edge.

[3.5.](#) Other benefits?

TBD: any other aspects to consider.

[3.6.](#) Security Considerations

TODO.

[4.](#) References

[4.1.](#) Normative References

[RFC7933] "Adaptive Video Streaming over Information-Centric Networking (ICN)", [RFC 7933](#), august 2016.

[4.2.](#) Informative References

[frecon1998dive]
and , "DIVE: A scaleable network architecture for distributed virtual environments", Distributed Systems Engineering vol 5, number 3 , 1998.

[funkhouser1995ring]
and , "RING: a client-server system for multi-user virtual environments", ACM symposium on Interactive 3D graphics , 1995.

[macedonia1995exploiting]
and , "Exploiting reality with multicast groups: a network architecture for large-scale virtual environments", Virtual Reality Annual International Symposium , 1995.

[su2013benefit]
and , "On the Benefit of Information Centric Networks for Traffic Engineering", IEEE ICC , 2014.

[van2010survey]
and , "A survey of augmented reality technologies, applications and limitations", International Journal of Virtual Reality , 2010.

Author's Address

Cedric Westphal
Huawei

Email: Cedric.Westphal@huawei.com