

COIN
Internet-Draft
Intended status: Informational
Expires: April 22, 2019

M. Montpetit
Triangle Video
October 19, 2018

In Network Computing Enablers for Extended Reality
draft-montpetit-coin-xr-00

Abstract

Augmented Reality (AR) and Virtual Reality (VR), combined as Extended Reality or XR, challenge networking technologies and protocols because they combine the features of fast information display, image processing, computing and forwarding. This document presents some of these challenges and how adding computing in the network could respond to them.

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 April 22, 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.

Table of Contents

1.	Introduction	2
1.1.	Requirements Language	3
2.	Definitions	3
3.	Extended Reality and In-Network Computing	4
3.1.	XR Network Requirements	4
3.2.	In-Network Computing Advantages in XR	5
4.	Enabling Technologies	6
4.1.	Information Centric Networking (ICN) and Named Data Networking (NDN)	7
4.2.	Network Coding	7
4.3.	Blockchains and Distributed Trust	8
5.	Conclusion	9
6.	Acknowledgements	9
7.	References	9
7.1.	Normative References	9
7.2.	Informative References	9
	Author's Address	10

[1.](#) Introduction

Virtual Reality (VR) and Augmented Reality (AR) taken together as Extended Reality or XR are at the center of a number of technological advances in many different fields, including not only gaming and entertainment but immersive journalism, remote diagnosis and maintenance, telemedicine, manufacturing and assembly and seat cities . They all share a number of stringent delay and bandwidth requirements to prevent confusing the brain whenever information about the virtual environment is not wholly consistent causing motion sickness symptoms [[VRSICK](#)]. Hence to now XR has been delivered mostly locally via combinations of computers and headsets with some cloud implementations being limited to time invariant imaging in one direction.

But with the emergence of the edge and the programmability of network elements all the way from the data center to the users the possibility of creating networked, multiparty/multisource and interacting XR comes closer to reality. This document wants to review what is necessary for the current localized and cloud supported XR to evolve to a more distributed and edge centric architecture to support advanced immersive application and services. It assumes that network programmability will enable to tailor the network to the XR requirements. This document is about requirements not solutions per se but will mention work that has already been done

towards a more networked XR including Information Centric architectures, Artificial Intelligence and in network coding. The networked functionality should enable to supplement local XR services and devices while keeping the very low latency and the very high data rates that are required by XR.

This document is intended as informative to both the networking and application research community. It does not address a specific network layer or protocol but provides architecture and system level specifications and guidelines. For example:

Latency: the physical distance between the XR content cloud of AR/VR and users are short enough to limit the propagation delay to the 20 ms usually cited for XR applications [ref] mixed for example with IoT devices and sensors delay reduction for range of interest (RoI) detection.

Applications: better coding and use of compression algorithms, pre-fetching and pre-caching and movement prediction.

Network access: push some networking functions in the data plane into the user plane to enable the deployment of stream specific algorithms for congestion control and application-based load balancing based on machine learning and user data patterns.

Network access: push some networking functions in the data plane into the user plane to enable the deployment of stream specific algorithms for congestion control and application-based load balancing based on machine learning and user data patterns.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)].

2. Definitions

AR: Augmented Reality (AR) is a live direct or indirect view of a physical, real-world environment whose elements are augmented by computer-generated input such as sound, video, location or graphical data. It is related to a more general concept called mediated reality [[MEDIA](#)], in which reality is modified (diminished or augmented) by computer-generated imagery.

VR Virtual Reality (VR): uses software-generated realistic imaging, sounds and other sensor inputs to replicate a real or imaginary setting, to simulates a user's physical presence in this

environment and provide an immersive experience that enable the user to interact with objects and move within this space.

360-degree video: 360-degree videos, also known as immersive videos or spherical videos, are video recordings where a view in every direction is recorded at the same time using an omnidirectional camera or a collection of cameras. 360o video is outside the scope of this document.

XR: extended reality is used to address both AR and VR together.

3. Extended Reality and In-Network Computing

XR is an example of the Multisource Multidestination Problem that combines video, haptics and tactile experiences as well, in interactive or networked mode multiparty and social interactions. Thus, XR is difficult to deliver to deliver with a client-server strictly cloud-based solution as it requires a combination of stream synchronization, lows delay and delay variations as mentioned above as well as means to cover from losses and provide optimized caching in the cloud and rendering as close as possible to the user at the network edge.

3.1. XR Network Requirements

In order to deliver the XR experience, there is a need to achieve complete 6 degrees of freedom meaning the 3 axes for body movement (x,y,z) plus pitch, yaw, rotation of the head all of which must be fulfilled in real time again focusing on the low delay, low loss and low delay variation to avoid sea sickness symptoms if the image does not follow the movement [[CABLE](#)]. But this is not the only difficulty, as there is also the need to provide real-time interactivity for immersive sports, mobile immersive applications with tactile and time-sensitive data and high bandwidth for high resolution images. Since XR deals with personal information and potentially protected content (in entertainment and gaming) XR must also provide a secure environment and ensure user privacy. And of course, the sheer amount data needed for and generated by the XR applications will use recent trend analysis and mechanisms, including machine learning to find these trends and reduce the size of the data sets.

Shared and global immersive experiences require interconnected, distributed and federated XR nodes. The requirements can be summarized as:

- Allow joint collaboration in VR.

- Provide multi-view AR.
- Add extra streams (IoT) to AR and VR experiences across services.
- Provide "Social Television" experiences and global viewing and experience rooms.
- Enable multistream, multidevice, multidestination applications.
- Use new Internet Architectures at the edge for improved performance.
- Integrate with holography, 3D displays and image processing systems [[CABLE](#)].

[3.2.](#) In-Network Computing Advantages in XR

One aspect of the push of XR to the edge is of course to provide cloud-based services with much lower latency. While this is very promising the question of the localization of the networking resources in order to provide the service becomes an essential component of the overall architecture. But it is not only finding the best geographical location but also providing the right level of reliability when one or more location is not available especially for mission critical services in medicine or manufacturing. And it does not mean only data laid distribution but also ensuring the availability of the right computational capabilities. The optimization of the location and type of the required resources for the multisource, multidestination, multiparty, multi-input XR applications can use AI and ML, and advanced load balancing and distributed network principles. There is a need for more research in such resource allocation problems at the edge to enable autonomous node operation and quality of experience [[SOL](#)]. These are of course multi-variate and heterogeneous goal optimization problems requiring advanced analysis with fast converging algorithms [[MULTI](#)][[PACKET](#)]. This is essential for the federation of nodes to provide the required experience.

Of course, image rendering and video processing in XR leverages different HW capabilities combinations of CPU and GPU. Current programmable network entities need to be evaluated to see if they can be sufficient to provide the speed required to provide real-time rendering and execute complex analytics: P4 for example does not support the floating-point operations necessary for advanced graphics.

Finally, dynamic network programmability could enable the use of joint learning algorithms across both data center, edge computers and goggle or glasses to allocate functionality and the creation of semi permanent datasets and analytics for usage trending. In the end, the use of computing or networked XR will enable the allocation of control, forwarding and storage resources and related usage models when needed by the application. This may mean re-evaluating the distribution of functionalities between datacenter and edge with less critical elements rendered in the cloud combined with a better understanding of the operational decomposition of the XR experience to allow the use of novel data structures, three-dimensional modeling and image processing algorithms.

Other advantages of adding computing to networked XR include:

- Multicast distribution and processing as well as peer to peer distribution in bandwidth constrained environments.
- Evaluation of local caching and micro datacenters with local or cloud-based pre-rendering.
- Trend or ML based congestion control to manage XR sessions quality of service.
- Higher layer protocols optimization to reduce latency.
- Trust, including blockchains and smart-contracts to enable secure community building across domains.
- Support for nomadicity and mobility (link to mobile edge).
- Use of 5G slicing to create independent session-driven processing/rendering.
- Performance optimization by tunneling, session virtualization and loss protection.

4. Enabling Technologies

This section presents some salient research that will lead to in-network computing becoming a major enabler of networked XR.

NOTE: more information and added sub-sections will be added in future versions of the draft with the collaboration of co-authors in the specific research areas.

4.1. Information Centric Networking (ICN) and Named Data Networking (NDN)

The Named Data Networking (NDN) architecture, one architecture of ICN, is particularly well suited for the multisource multi-destination architecture of XR because it allows to create the content experiences based on their components names not a location or pointer to a location hence provides a natural functional decomposition. ICN allows content delivery to evolve from single, context-independent streams to context-dependent Information components that can adapt dynamically to the changes necessary to maintain the immersive nature of the experience and be delivered efficiently. The combination of interest messages to signal what content is needed combined with the data responses help to coordinate the different streams and multiple users (pull mechanisms). The ICE-AR [[ICE](#)] project already mentions a concept of acceleration as a service: the exploration of the design and the usage of computation at the edge including the wireless edge.

For XR, ICN also allows to develop robust and resilient networking while allowing application developer to continue using known programming model [[RICE](#)]. This is important for the XR developers community that come from the entertainment, gaming or other non network specific industries and could enable ICN and XR to coexist in user devices (the ultimate edge). NDN concepts are already integrated to distributed video distribution with trust mechanisms (see section below) such as smart contracts on the blockchain to proof of origin and destination sent along with interest messages [[HUITX](#)].

4.2. Network Coding

Networked XR requires the synchronization of multiple streams but with its delay sensitivity the use of buffering schemes to achieve this synchronization is impractical. At the same time the need to maintain high image quality means that packet losses also need to be limited. Network coding has proven very useful to achieve both these goals in commercial streaming services like Netflix, is being added to protocols like QUIC and in another multi-stream service namely Social Television [[SOCIAL](#)] avoiding the reliance on complex synchronization algorithms. The main difference between XR and Social Television is that the former is even more constrained in latency and loss budgets hence even the delay due to encoding and decoding operations needs to be minimized. Hence the idea of in-network coding and re-encoding to adapt to dynamic network conditions, not just end to end, can be used to ensure on time packet delivery with loss recovery. In network encoding needs the type of programmability that COIN provides.

4.3. Blockchains and Distributed Trust

If XR is to be integrated at the edge of the network to provide the required delay and loss guarantees, then relying on centralized mechanisms for trust is non-realistic. Traditional centralized mechanisms to discover and admit nodes to the network, to provide access right and name resolution need to be updated to be used in the dynamic XR environment. Blockchain technology, with operation performed at the edge and in a decentralized way is fast becoming a major scalable means of providing trust and validate provenance in a large number of applications including those on the XR portfolio. Smart contracts (on the blockchain) supply a mechanism to provide the trust and validation for XR edge nodes.

A new XR participant node is admitted after it has committed to a smart contract that contains the rules and mechanisms to distribute content via this node in a trusted and secure way. This constitutes its proof of validity. After a node is admitted, it will can then provisioned with the appropriate software to become fully operational to provide the XR experience. Newly admitted nodes will be inserted in the general ledger on the blockchain enabling other nodes to discover them, and hence, to form a trusted network. A name resolution authority can also be provided by the blockchain to manage and validate the origin of the content, the proof of origin, and to provide the ability to search such content. The proof of origin can also be used to prevent some content from reaching one or more nodes and implement content filtering based on trusted authorities. This is useful not only for content packets but also for packets capable of modifying the node operations. Finally, when some content reaches a specific destination, it can be verified against the content rules of the reached node even and before it is sent to the application; this allows to provide a proof of delivery for the content and enable to generate statistics, performance metrics and enable the nodes to adapt to the XR requirements.

All of the above assumes that the nodes can implement the functions needed by the blockchain hence once again infers that there is enough computing power in the nodes to perform these operations. At this point both proof of concept and proof of every are limited due to the added overhead and the size of the blockchain. As distributed blockchain and COIN continue to evolve this should continue to be a field of interest for the development of secure and private XR experiences.

5. Conclusion

More and more applications and service are being developed and deployed that use or will use combinations of AR and VR, XR, along with extra stream from sensors and IoT devices. And many of these applications require to be deployed over a network because of their interactive or multiparty nature. In that context, it not uniquely necessary to move functionality to the network but to carefully evaluate which elements to locate in network nodes, where these nodes are and what computational support they need to support the XR experience. Hence, it is believed that a great enabler of networked XR is the capability to co-locate programmable elements in the XR network node to respond to the dynamics of the services in an efficient, resilient and secure manner.

6. Acknowledgements

The author would like to thank Jeffrey He, Dirk Kutscher, Cedric Westphal and Weiguang Wang for their contributions to the presentation that lead to this draft.

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.

7.2. Informative References

[CABLE] Hinds, A., "The Near Future of Immersive Experiences: Where We Are on the Journey, What Lies Ahead, and What It Takes to Get There", SIGCOMM 2018 Workshop on AR/VR <http://conferences.sigcomm.org/sigcomm/2018/workshop-arvr.html>, August 2018.

[HUITX] "8X: ICN Based Video Distribution", 2018, <<https://www.8xlabs.com>>.

[ICE] Burke., J., "ICN-Enabled Secure Edge Networking with Augmented Reality: ICE-AR", ICE-AR Presentation at NDNCOM September 2018 <https://www.nist.gov/news-events/events/2018/09/named-data-networking-community-meeting-2018>, 2018, <<http://ice-ar.named-data.net>>.

- [INTER] Bastug et al., E., "Towards Interconnected Virtual Reality: Opportunities, Challenges and Enablers", IEEE Communications Magazine, Volume 55 , Issue: 6 , 2017 <https://arxiv.org/pdf/1611.05356.pdf>, June 2017.
- [MEDIA] "Mediated Reality", Wikipedia.org https://en.wikipedia.org/wiki/Computer-mediated_reality, 2018.
- [MULTI] Batalla, J., "Evolutionary Multiobjective optimization algorithm for multimedia delivery in critical applications through Content-Aware Networks", The Journal of Supercomputing, Volume 73, Issue 3, pp. 993-1016 <https://link.springer.com/article/10.1007/s11227-016-1731-x>, March 2017.
- [PACKET] Jeyakumar et al., V., "Millions of Little Minions: Using Packets for Low Latency Network Programming and Visibility", Proceedings of SIGCOMM 2014 <http://conferences.sigcomm.org/sigcomm/2014/program.php>, August 2018.
- [RICE] Krol et al., M., "RICE: Remote Method Invocation in ICN", Proceedings of the ACM Conference on Information-Centric Networking 2018 <http://conferences.sigcomm.org/acm-icn/2018/proceedings/icn18-final9.pdf>, September 2018.
- [SOCIAL] Montpetit, M. and M. Medard, "Social Television: Enabling Technologies and Architectures", Proceedings of the IEEE, Volume 100, pp. 1395-1399 <http://proceedingsoftheieee.ieee.org>, May 2012.
- [SOL] Heorhiadi et al., V., "Simplifying Software-Defined Network Optimization Using SOL", 13th USENIX Symposium on Networked Systems Design and Implementation <https://www.usenix.org/system/files/conference/nsdi16/nsdi16-paper-heorhiadi.pdf>, March 2016.
- [VRSICK] LaViola, J., "A Discussion of Cybersickness in Virtual Environments", ACM SIGCHI Bulletin 32(1):47-56 <http://www.eecs.ucf.edu/~jjl/pubs/cybersick.pdf>, January 2000.

Marie-Jose Montpetit
Triangle Video
Boston, MA
US

Email: marie@mjmontpetit.com