Will Metaverse Be NextG Internet? Vision, Hype, and Reality

Ruizhi Cheng, Nan Wu, Songqing Chen, and Bo Han

Abstract

Metaverse, with the combination of the prefix "meta" (meaning transcending) and the word "universe," has been deemed as a hypothetical next-generation (NextG) Internet. It aims to create a shared virtual space that connects all virtual worlds via the Internet, where users, represented as digital avatars, can communicate and collaborate as if they are in the physical world. Nevertheless, there is still no unified definition of the Metaverse. This article first reviews what has been heavily advocated by the industry and the positions of various high-tech companies. It then presents our vision of what the key requirements of Metaverse should be. After that, it briefly introduces existing social virtual reality (VR) platforms that can be viewed as early prototypes of the Metaverse, and conducts a reality check by diving into the network operation and performance of two representative platforms: Workrooms from Meta and AltspaceVR from Microsoft. Finally, it concludes by discussing several opportunities and future directions for further innovation.

Introduction

Although the term Metaverse has been around for almost 30 years since it was coined by Neal Stephenson in his 1992 science fiction novel Snow Crash, we are still in the early stage of actually building the Metaverse, which envisions an immersive successor to the Internet. The development of the Metaverse has gone through several stages. Retrospectively, the text-based interactive games, such as MUD (multi-user dungeon) that emerged in the late 1970s and defined a multiplayer virtual world with role playing, interactive fiction, and online chat, could be viewed as the earliest prototypes of Metaverse, even before the term was literally introduced. The second phase happened during the postmillennial decade with the development of commercial virtual worlds such as Second Life (https://secondlife.com/; accessed on 25-Aug.-22). It then embraced fully 3D virtual worlds such as OpenSimulator (http://opensimulator.org/; accessed on 25-Aug.-22), which is largely compatible with Second Life.

In the current stage, with the flourishing of 5G and mobile immersive computing [1], there has been a surge of research & development on the Metaverse in both industry and academia. We have now entered an open development phase of the Metaverse, which is widely considered as a collection of 3D virtual worlds connected via the Internet [2] and enabled by various immersive technologies such as augmented reality (AR), virtual reality (VR), and mixed reality (MR), which are often collectively referred to as extended reality (XR). While there is still no unified definition of the Metaverse, it is broadly deemed as a hypothetical next-generation (NextG) Internet (https://bit.ly/3cn5SCr; accessed on 25-Aug.-22).

Figure 1 illustrates the basic elements in the Metaverse and how they interact with the physical world. In general, users with XR devices access the Metaverse and participate in its various social events, whose smooth execution is enabled by techniques such as 5G and HCI (human-computer interaction). Users are free to create their own content via 3D modeling for decorating social events in the Metaverse. The content can be traded using non-fungible tokens (NFTs) through a decentralized blockchain. Physical objects can be presented in the Metaverse as digital twins that are generated via 3D modeling and consumed with XR devices assisted by artificial intelligence (AI).

In this article, we first review recent advances in the industry and introduce the advocates of various key players. We then present our vision of the Metaverse by discussing its key technical requirements. After that, we provide an overview of existing social VR platforms, the early prototype of Metaverse that combines online social networks and VR technologies, and compare their unique features. We then conduct a first-of-its-kind reality check to understand the networking protocol usage and system performance of two representative platforms, Meta's Horizon Workrooms (https://www.oculus.com/workrooms/; accessed on 25-Aug.-22) (referred to as Workrooms) and Microsoft's AltspaceVR (https://altvr.com/; accessed on 25-Aug.-22). Finally, we discuss the technical challenges, opportunities, and directions for future research activities and conclude this article.

Industry Trends

In this section, we briefly introduce the current development of Metaverse in the industry.

As shown in Fig. 2, many high-tech companies have joined the Metaverse arena. Meta is conceivably the most notable among all that have invested in this space. In September 2019, Meta (named Facebook then) announced *Facebook Horizon*, a social VR platform. In July 2021, Facebook announced the transition into a Metaverse company within five years. To echo this vision, in October 2021, Facebook changed its name to

The authors are with George Mason University.

Digital Object Identifier: 10.1109/MNET.117.2200055

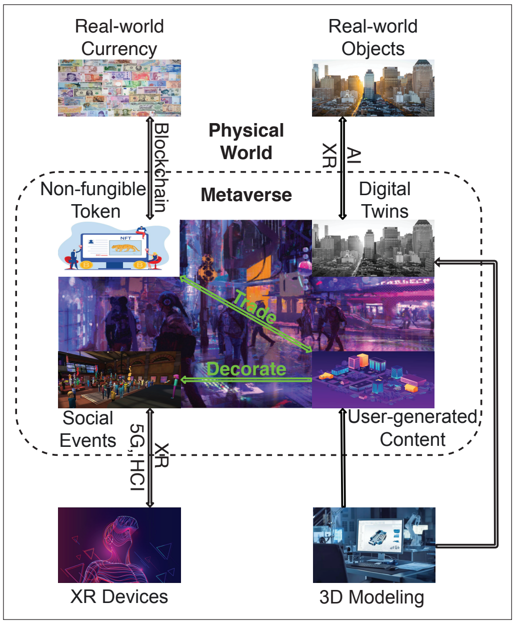

FIGURE 1. Elements of the Metaverse and their interaction with the physical world.

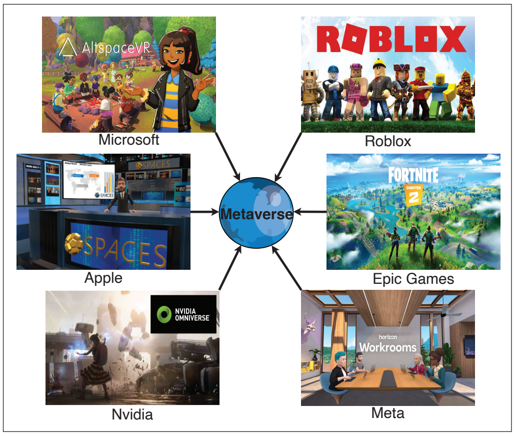

FIGURE 2. Current development of the Metaverse in industry.

Meta. Meta considers VR to be the foundation to build the Metaverse. Its VR headset, Oculus Quest 2, has sold over 10 million units (https://bit.ly/3pHJXsY; accessed on 25-Aug.-22), making it the best-selling VR device.

Nvidia announced a plan to create the first virtual collaboration and simulation platform called Omniverse (https://www.nvidia.com/en-us/omniverse/; accessed on 25-Aug.-22) in August 2021. This platform can be used to connect 3D worlds into a shared virtual universe and create digital twins, simulating real-world buildings and factories. Omniverse has three key components. The first one is Omniverse Nucleus, a database engine that allows multiple users to connect and create a scene together. The second one is the rendering and animation engine to simulate the virtual world. The third one is Nvidia CloudXR for streaming XR content to client devices. Meanwhile, Omniverse integrates AI to train digital twins in the Metaverse.

Epic Games, a video game company famous for its Unreal game engine, announced a $1 billion investment to build the Metaverse. In its most popular game, Fortnite (https://www.epicgames.com/fortnite/en-US/home; accessed on 25-Aug.-22), users can create their avatars, buy digital items, and enjoy movies and concerts. Roblox (https://www.roblox.com/; accessed on 25-Aug.-22) is another company in this arena. As the largest UGC game platform, players in Roblox can create their own games and virtual worlds. They can buy, sell, and create virtual items that can be used to decorate their avatars.

Although most companies embrace the Metaverse's concepts and vision, cautions and doubts also emerge. While both Apple and Microsoft have virtual space applications,[1] they consider that seamlessly connecting the Metaverse and the physical world is a key to its success, if not more important than the Metaverse itself. They believe that the purpose of creating the virtual space is just to enable users to improve productivity and reduce production costs in the physical world.

Defining Metaverse

Existing Definitions and Enabling Technologies

Metaverse has been viewed as a new type of online social network, and arguably NextG Internet. It would be the convergence of digital second life (for "escape") and virtual reality (for exploration), mimicking user interaction in the real world. A narrow definition of Metaverse is thus a universal virtual world focusing on social interaction, which connects multiple 3D virtual environments via the Internet (i.e., a network of virtual worlds [2]). We envision that the Metaverse should evolve to *seamlessly integrate the physical world and the virtual space*, for example, via digital twin and digital economies (e.g., cryptocurrencies).

Objects in the physical world can interact with the Metaverse. They can generate their digital twins through 3D modeling and keep their digital twins presenting the same state as what is happening in the real world. Conversely, after the digital twin is manipulated/processed in the Metaverse, its physical-world state will be changed accordingly. For example, BMW has used Omniverse from Nvidia to construct a fully functional digital twin of its automobile factory, reducing manufacturing costs and increasing productivity.

While there is no consensus on the definition, as shown in Fig. 3, it is commonly agreed that the Metaverse is built on and integrates technologies such as 5G, XR, edge computing, blockchain, machine learning (ML), and HCI [3].

5G provides a faster, lower latency, and more scalable network than 4G. According to the frequency bands, 5G can be divided into low-band (below 1 GHz), mid-band (between 1 and 6 GHz), and high-band (millimeter-wave, mmWave, from 24 to 39 GHz). Low-band 5G is used for extensive coverage and is ideal for deployment in rural areas. Mid-band 5G has been commonly deployed in metropolitan areas. High-band 5G can reach a maximum throughput of, in theory, 10-20 Gb/s. However, it works in only a small radius, and thus is more useful in urban areas and crowded locations (e.g., shopping malls).

AR/VR/MR augment or supplant our view of the world, and are a key to the success of Metaverse [3]. VR immerses people in the virtual world, and social VR is widely considered an important component of the Metaverse. AR

[1] Apple acquired a VR company, Spaces, in 2020, and Microsoft acquired a social VR platform called Altspace-VR back in 2017.

enables digital twins in the Metaverse to be overlaid on physical objects in a perceptible way, effectively connecting the Metaverse with the physical world. MR allows users to interact with virtual objects, by creating more connections and collaborative relationships among the physical world, virtual space, and users.

Edge Computing is a computing paradigm that moves computation and data storage closer to users. The advantageous performance of edge computing in reducing latency for XR has made it an important backbone for building the Metaverse. Several telecom carriers have undertaken a project called HoloVerse to test the best 5G edge network infrastructure for efficient deployment of services in the Metaverse (https://bit.ly/3wtYjki; accessed on 25-Aug.-22). Meanwhile, Niantic, the producer of *Pokémon-Go*, has joined forces with telecom carriers to explore how 5G edge computing can enhance the quality of experience (QoE) for AR games (https://bit.ly/3L1Uj0r; accessed on 25-Aug.-22).

Blockchain ensures the security of data records and generates trust without requiring trusted third parties. It is closely related to user-generated content (UGC) such as digital assets that can greatly enrich the Metaverse [4]. For example, NFT, which is used for trading in the Metaverse, is a data unit on the blockchain. Defining the ownership of UGC in the Metaverse is a practical challenge, as digital assets can be copied and reproduced. NFT provides an effective way to prove that UGC is unique and non-fungible (i.e., non-interchangeable). It enables owners of digital content to sell/trade their property via smart contracts in the decentralized crypto space (e.g., using cryptocurrencies).

Machine Learning, especially deep learning (DL), is an important branch of artificial intelligence (AI) that enables machines to learn from massive amounts of data. Undoubtedly, the Metaverse will generate a huge amount of complex data, providing rich opportunities for DL. For example, we can leverage DL to design digital twins, which can autonomously collect data from the Metaverse and the physical world in real-time for model training and inference.

HCI enables users to interact with digital twins in the Metaverse in real-time. One of the most important HCI problems to be addressed is user input. The key limitation of existing input devices (e.g., mice and keyboards) is that they cannot free the users' hands and accurately reflect their body movements. Recently, researchers have begun to study freehand manipulation that allows more intuitive and concrete interaction in the Metaverse. These techniques often rely on computer vision and brain-computer interfaces.

Our Vision on Technical Requirements

Next, we present our vision of the Metaverse by illustrating three key requirements on scalability, accessibility, and security, privacy, and legal issues.

Requirement #1: Scalability: With the Internet transitioning to the Metaverse, we expect the first practical challenge faced by any Metaverse platform is the scalability issue. As our preliminary investigation below shows, current social VR platforms such as Workrooms and AltspaceVR, an early prototype of the Metaverse, can hardly scale up to tens of participants. When more participants access Workrooms, the corresponding uploading and downloading demand increases proportionally. The platform, either serving just as a relay or performing further content processing in the middle, will quickly become a bottleneck.

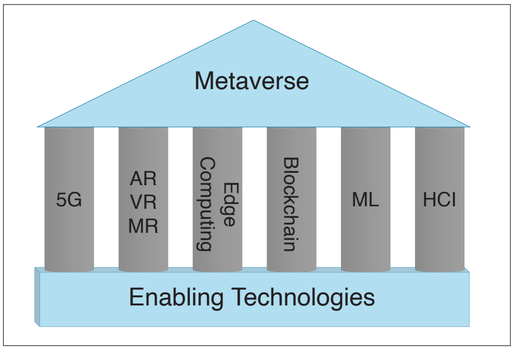

FIGURE 3. Enabling Technologies of the Metaverse.

As can be expected, the bandwidth requirement of the Metaverse could be huge. On one hand, compared to traditional 2D videos, the bandwidth for transmitting up to 16K 360-degree panorama [5] or 3D volumetric content [6] to XR headsets could be high. On the other hand, the Metaverse is full of social elements, which further increases the bandwidth requirement. Currently, the U.S. Federal Communications Commission (FCC) defines the standard broadband service as 25 Mb/s in downlink and 3 Mb/s in uplink [7]. Therefore, it is of the utmost importance to guarantee the scalability of Metaverse by leveraging advanced networking techniques.

Requirement #2: Accessibility: Today's Internet access does not need specific devices. For the Metaverse, however, users are required to wear headsets for better interaction in the virtual world. It greatly limits the accessibility of the Metaverse, mainly due to the inconvenience of such access. We envision that in the future, new "interfacing" devices should be developed for accessing the Metaverse without wearing any additional device, and glasses or contact lenses would replace the cumbersome headsets [3]. Moreover, interaction techniques, other than just display, would need to be in place so that users can not only see in the virtual world but also feel, smell, taste, and so on, like what we do in the physical world [2].

Besides the interfacing devices of the Metaverse, another potential obstacle is network accessibility. The average 25 Mb/s downlink bandwidth in the U.S. [7] is far from the demand of even a rudimentary Metaverse — the bandwidth requirement would go up with more and more user-generated content and assets in the Metaverse. Yet another issue related to accessibility is the interoperability across different implementations of the Metaverse, especially when users move from one platform to another. The user experience should be seamless without any interruption.

Requirement #3: Security, Privacy, and Legal Issues: Similar to online social networks, in the Metaverse, there will be security and privacy issues, such as attacks on user authentication and impersonation [8]. Meanwhile, users' personal information (e.g., biometric data) may be collected for authentication, compromising their

Platforms	Company	Quest 2	Facial Expression	Game Events	Personal Space	Share Screen	Shopping	NFT
AltspaceVR ('15)	Microsoft	✓	×	✓	✓	✓	×	×
Bigscreen ('16)	BigScreen	✓	×	×	×	✓	×	×
Rec Room ('16)	Rec Room	✓	✓	✓	✓	×	✓	✓
Anyland ('16)	Anyland	×	×	✓	×	×	×	×
VRChat ('17)	VRChat	✓	✓	✓	✓	×	×	×
Cluster ('17)	Cluster	✓	✓	✓	×	×	×	×
Hubs ('18)	Mozilla	✓	×	×	×	✓	×	×
Workrooms ('21)	Meta	✓	✓	×	×	✓	×	×

TABLE 1. Comparison of several important features offered by eight social VR platforms. Personal space is a protective zone in the virtual environment that users can define.

privacy [9]. Moreover, there will be new types of challenges, for example, securing the NFTs and digital twins, which involve interaction with the physical world. Furthermore, online harassment can be exacerbated in the immersive environment of the Metaverse by features including free avatar movements and enhanced feelings of presence and embodiment [10]. Additionally, given that the Metaverse assets (content) are user-generated, there will be copyright issues. The protection of content ownership, the detection of copyright infringement, and the licensing of such content have not been well laid out. Considering that there will be multiple Metaverse platforms, transferring users' assets from one to another is a practical issue to be addressed. Such portability and interoperability demand not only standardizations from the industry but also legal enforcement.

Social VR Platforms

Since social VR is a major component of the Metaverse, we provide an overview of several commercial social VR platforms, highlighting their key features and differences. Social VR, regarded as the future of social media, allows users to interact with each other as avatars in the virtual world, communicating and collaborating as if they are in the physical world. With the global outbreak of the COVID-19 pandemic, many people around the world have to stay at home and lack social interactions, leading to surging demand for novel applications of social media. Thus, predictably, the demand for social VR will continue to grow, as it not only satisfies people's social needs but also gives them a sense of spatial presence.

Key Features

After an extensive survey, we focus on the eight most popular social VR platforms, VRChat (https://hello.vrchat.com/; accessed on 25-Aug.-22), Rec Room (https://recroom.com/; accessed on 25-Aug.-22), AltspaceVR, Mozilla Hubs (https://hubs.mozilla.com/; accessed on 25-Aug.-22), Anyland (http://anyland.com/; accessed on 25-Aug.-22), Cluster (https://cluster.mu/en/; accessed on 25-Aug.-22), Bigscreen (https://www.bigscreenvr.com/; accessed on 25-Aug.-22), and Workrooms. As a first step, we mainly focus on examining the following questions:

- Whether they are accessible from the popular Oculus Quest 2
- Whether their avatars have facial expressions
- Whether they have the personal space feature, which is a zone to protect users from harassment
- Whether they have the gaming, sharing PC screen, and shopping features
- Whether they support the trading of assets with NFTs.

Table 1 presents a summary of these platforms. We find that all platforms except Anyland support Oculus Quest 2. Avatar's facial expression, game events, personal space, and PC screen sharing are supported by about only half of the platforms, showing varied design choices and development stages across them. Finally, only Rec Room offers shopping and NFTs, demonstrating that virtual trading is not yet widely available on social VR platforms.

User Experience

We experiment with the above platforms and highlight their advantages in terms of QoE.

AltspaceVR: The ambient lighting of the virtual scene matches the shadows, making the lighting of the scene realistic. There are many environments and live events initiated from all over the world, with a rich social element.

Bigscreen: Users can play PC games in the virtual world and watch together videos (e.g., Netflix and YouTube) played on PCs in a private or public room.

Rec Room: It offers an abundance of game activities and enables cross-play between different users with VR headsets, PCs, and smartphones.

Anyland: It is a "sandbox universe," where users can build anything (even the avatar) they need using tools that exist in the physical world.

VRChat: Users can build their own games in the virtual world. It allows an impressive amount of customization (e.g., users can upload any 3D model as the avatar).

Cluster: It has a variety of highly interactive social events, such as live concerts. In addition, it can make souvenir books with the photos taken by users at each event.

Mozilla Hubs: Users can customize their applications with its source code and deploy their own servers. They can use Hubs through browsers without downloading any application, which is lightweight and convenient.

Workrooms: It supports physical keyboards, which are much more convenient than the virtual ones manipulated by a controller. Moreover, users can write using the controller as a pen by flipping it around.

Other Types of Metaverse Platforms

In addition to social VR, recent advancements of the Metaverse embody massively multiplayer online games, such as Fortnite, Minecraft (https://www.minecraft.net/en-us; accessed on 25-Aug.-22), and Roblox, as well as the emerging NFT or blockchain-based online games, such as Decentraland (https://decentraland.org/; accessed on 25-Aug.-22), Upland (https://upland.me/; accessed on 25-Aug.-22), and Axie Infinity (https://axieinfinity.com/; accessed on 25-Aug.-22). However, since they are designed primarily for PC users with 2D content, these games currently cannot provide their users with an immersive experience, which is one of the most important goals of the Metaverse.

Hence, in this article, we focus on the investigation of social VR platforms.

Case Studies

In this section, we conduct a reality check of the Metaverse by comparing the network operation and performance of Workrooms and AltspaceVR. As shown in Table 1, AltspaceVR is one of the earliest initiatives, and Workrooms is the most recent effort of the Metaverse, both supporting the majority of the listed features. Thus, comparing them will give us a better understanding of the state-of-the-art of Metaverse.

In our previous work [11], we dissected how Workrooms works. Our key findings are as follows:

- Workrooms primarily employs two servers to communicate with its clients, one is for delivering virtual content and the other is for streaming/exchanging audio and video data, as shown in Fig. 4 (top).
- Workrooms requires ~25s to initialize, by primarily performing local setup and rendering without much network activity (Fig. 5a and b).
- With two headset users in Workrooms, each user's downlink throughput is about 2–3 Mb/s and the uplink throughput is ~0.6 Mb/s (Fig. 5a). However, the downlink throughput linearly increases with the number of headset users, indicating that the current design of Workrooms may face scalability issues (Fig. 5b).
- Workrooms does not consider situations not requiring server involvement, but simply lets the server process and forward all users' data, resulting in unnecessary communication overhead (Fig. 5c).

We perform the same experiments on AltspaceVR to understand the differences between the two platforms. We conduct a series of experiments with a 3-minute duration. We use a Macbook Pro as the WiFi access point, which is connected to a high-speed home network via Ethernet for Internet access. We capture and analyze network traffic using the Wireshark packet analyzer (https://www.wireshark.org/; accessed on 25-Aug.-22).

Network Protocol Analysis

We first compare the network protocols employed by Workrooms and AltspaceVR. Besides headsets, we use Google Chrome to access Workrooms and the Windows application to access AltspaceVR from a PC. We find that users' devices communicate with two servers in both Workrooms and AltspaceVR. Figure 4 summarizes the process of establishing connections and exchanging data between the clients and the servers in Workrooms (top) and AltspaceVR (bottom).

In Workrooms, the connection with Server I starts during the loading period (i.e., when the loading progress bar is displayed). All data exchanges are over User Datagram Protocol (UDP). We have proven that this flow transmits virtual content [11] and refer to it as virtual content (VC) flow. The connection with Server II starts when users enter the meeting room. The headset and browser clients have a slightly different way of establishing connections with Server II. First, they both establish a Transmission Control Protocol (TCP) connection with Server II, while using Session Traversal Utilities for NAT (STUN) protocol

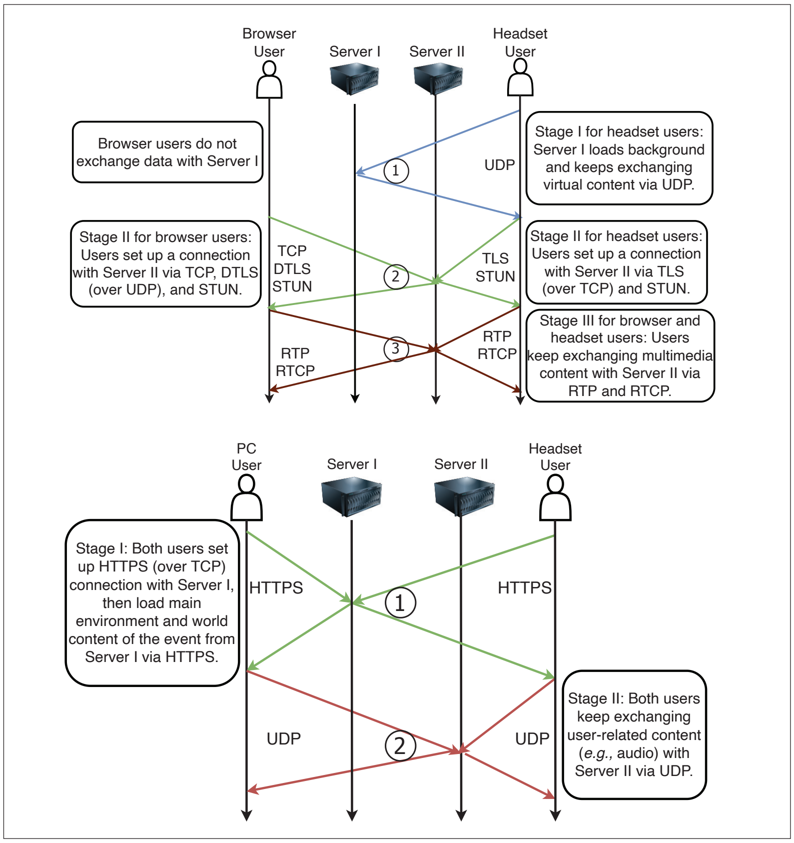

FIGURE 4. The process of establishing connections and exchanging data between the clients and the servers for Workrooms (top) and AltspaceVR (bottom).

to traverse network address translator (NAT) gateways. The headset and Server II then transfer 1-3 Transport Layer Security (TLS, a secure communication protocol over TCP) packets to each other. However, the browser client establishes a Datagram Transport Layer Security (DTLS, a secure communication protocol over UDP) connection with Server II. Finally, both browser and headset clients use Real-time Transport Protocol (RTP) and RTP Control Protocol (RTCP) to exchange multimedia content with Server II. We refer to this flow as multimedia (MM) flow.

In AltspaceVR, however, the client-server connections work in a different way. First, the client downloads 10-20MB of data from Server I using the Hypertext Transfer Protocol Secure (HTTPS) protocol, when the client screen displays "downloading world content." Only users who join the event for the first time need to download the data. Then, the client downloads 300-500KB of data from Server I via another HTTPS connection, when the client screen displays "loading main environment."

The connection with Server II starts when users finish loading. All data exchanges are over UDP. Since this UDP flow is the only flow after users enter the event, our hypothesis is that it contains user-related content. Through further analysis, we find that this UDP flow follows a custom protocol. The fourth byte of the UDP payload is used to distinguish the data type, such as audio data.

Network Performance

Next, we compare the network performance of two platforms based on key findings of Workrooms. The scalability experiments involve up to

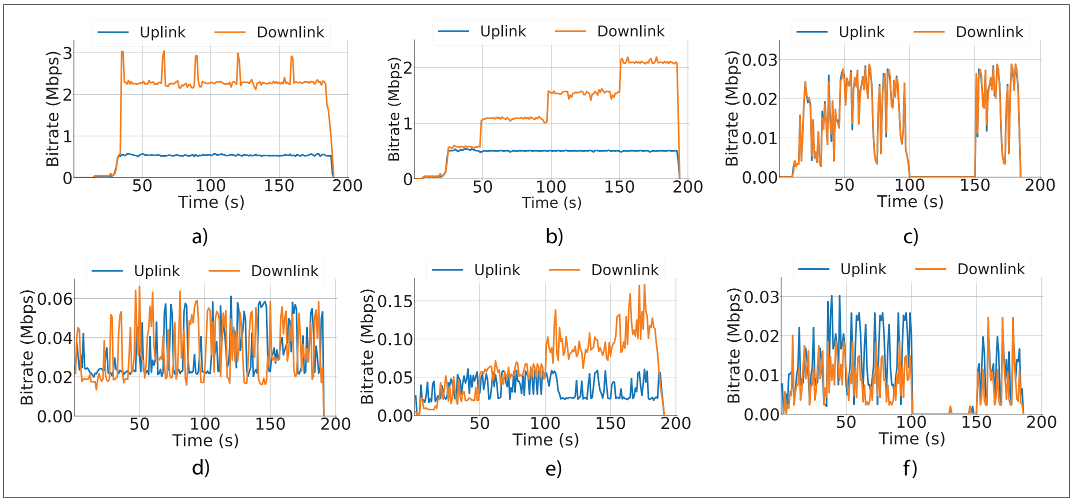

FIGURE 5. Comparison of throughput, scalability, and audio data between Workrooms and AltspaceVR. a) and d) U_1's uplink and downlink throughput of UDP flow in Workrooms and Altspace VR; b) and e) U_1's uplink and downlink throughput of VC flow (no change for MM flow) in Workrooms and UDP flow in AltspaceVR (three additional users U_3, U_4, and U_5 join at 50, 100, and 150 s, respectively); c) and f) comparison of audio data for U_1's uplink and U_2's downlink in Workrooms and AltspaceVR (both users mute from 100 to 150 s).

five users (denoted as U_i), all with Oculus Quest 2, and all other experiments have only two users (i.e., U_1 and U_2). Figure 5d shows the throughput (i.e., bitrate) of U1's UDP flow in AltspaceVR with two users. We find that since users have downloaded the content before entering the event, the throughput (less than 0.06 Mb/s) after that is much lower than Workrooms. Figure 5e shows the throughput of U_1's UDP flows in AltspaceVR, where we let three other headset users join the experiment at 50, 100, and 150 s, respectively. We find that AltspaceVR also faces scalability issues, with the downlink bitrate increasing almost linearly every time a new user joins (~0.03 Mb/s). However, its increase is much smaller than Workrooms (~0.5 Mb/s).

Figures 5c and f show the comparison of the audio data for U_1's uplink and U_2's downlink. In Workrooms, we observe that the bitrate on the downlink of U_1 largely matches that of the uplink of U_2, and vice versa (Fig. 5c). This indicates that the server simply forwards one user's audio data to others without further processing. However, in AltspaceVR, the uplink audio data of one user does not match the downlink audio data of another user, indicating that the server processes the audio data before forwarding it (Fig. 5f). Also, most of the time, the downlink throughput of a user is lower than the uplink throughput of the other user, which indicates that the server may optimize the audio data uploaded by users.

Accessibility and Security and Privacy Issues

Finally, we study the accessibility and potential security and privacy issues, considering that these are the key requirements to the success of the Metaverse.

AltspaceVR supports nearly all VR headsets, whereas Workrooms supports only the Oculus Quest 2 headset.

As shown in Table 1, neither AltspaceVR nor Workrooms offers shopping/NFTs. Hence, there are currently limited concerns regarding the security of transactions. However, with the active participation of high-tech companies such as Meta (https://bit.ly/3zc59gc; accessed on 25-Aug.-22) in the NFT, these concerns may emerge in the near future. In addition, AltspaceVR provides the personal space that protects users from harassment, but Workrooms does not.

To summarize, by comparing Workrooms and AltspaceVR, we have the following findings:
- AltspaceVR requires users to download event data in advance and consumes less bandwidth than Workrooms.
- Both platforms face potential scalability issues, although AltspaceVR does not cause significant bandwidth consumption (~0.18 Mb/s for downlink with five users).
- Unlike Workrooms, AltspaceVR processes the data uploaded by users before forwarding it, reducing the size of data received by other users.

Discussion

In this section, we discuss the challenges of building the Metaverse and point out opportunities for further innovation.

Technical Challenges

First, the operation of the Metaverse will generate a large amount of data, such as metadata created by sensors, a shared virtual space for users' social activities, and the transmission of high-resolution video streams, requiring a huge network bandwidth. However, the existing 5G technology may not be sufficient to support the Metaverse. As shown in Fig. 5, the throughput required for current social VR platforms is low. This is because the avatar of the majority of existing social VR platforms has only the upper torso, and the movement of the avatar is not driven by the actual movement of the user but operated by the hand-held controllers of the VR headset. However, the future Metaverse will necessitate high-quality full-body avatars to provide a truly immersive experience for millions of concurrent users in the shared virtual environment, demanding higher throughput than we have observed on today's social VR platforms. For example, existing work such as Holoportation [12] reveals that the bandwidth required to deliver a photo-realistic 3D model of a human body by capturing its motion in real-time can exceed 1 Gb/s. Moreover, our recent measurement study [13] demonstrates that the scalability issue identified earlier also exists on other popular commercial social VR platforms including Mozilla Hubs, Rec Room, VRChat, and Horizon Worlds (https://www.oculus.com/horizon-worlds/; accessed on 25-Aug.-22).

To this end, we discuss several potential solutions to address the scalability issue. Viewport-adaptive optimizations, which aim to deliver mainly the content visible to users for saving bandwidth, can alleviate the scalability issue. However, if enormous avatars are visible in the user's viewport, the required network bandwidth to transmit their data and resource utilization on the device for rendering may still be high. Moreover, the server needs to predict the future viewport of the user to determine the to-be-delivered content, which may negatively impact the user experience if the prediction is inaccurate [5]. Another potential direction is peer-to-peer (P2P) communication techniques. In P2P, user devices will need to combine the content received from multiple parties and then render the virtual world accordingly. However, given that rendering is still performed on the client side, the on-device resource consumption could be excessive.

Another promising strategy to address the aforementioned scalability issue is to utilize remote rendering [14], in which the server is responsible for performing rendering tasks. In this scenario, even though there are a significant number of concurrent users (especially when their avatars are clustered together), the servers will render the entire scene in a user's viewport into a 2D video frame. Hence, the amount of transmitted data is independent of the number of users, alleviating the scalability issues. Nevertheless, remote rendering still poses technical challenges. For instance, similar to viewport-adaptive optimizations, the performance of remote rendering depends on the accuracy of viewport prediction. Additionally, the server may have to render the same number of scenes as the number of users since different users may have different viewports.

Second, network latency is critical to the user experience. Given that users may access the Metaverse from different parts of the world, ensuring low latency when users are across geographically distributed regions is a practical challenge. Meanwhile, sensors in the Metaverse, such as those on XR headsets and haptic devices, require latency as low as tens of milliseconds to maintain an immersive user experience [15]. Similar to the motion-to-photon latency in VR, in the Metaverse the latency between the motion of a user and its reflection perceived by others is a key metric to optimize.

Third, the security and privacy issues in the Metaverse deserve our attention. Although commercial social VR platforms employ secure communication protocols (e.g., TLS and DTLS) to protect transmitted data, as verified in our measurement study, the Metaverse may still lead to many security concerns, such as users' identification information. Since it requires users to access with headsets, they often need to identify themselves with biometric information, which could be a target of security attacks [9]. Digital twins in the Metaverse also need proper protection. There will be a large number of complex ML models for supporting digital twins, which in turn influence objects in the physical world. If these models are attacked, there will be unpredictable consequences in the physical world. Storing digital twins of the Metaverse in the blockchain is a possible direction [4].

> Unlike Workrooms, AltspaceVR processes the data uploaded by users before forwarding it, reducing the size of data received by other users.

Besides data privacy, harassment is another emerging concern in the Metaverse. The Metaverse has not only text-based or voice-based harassment, which has been studied for traditional social media, but also body movement-based harassment that can reflect the users' movement via their avatar through various sensors. However, the protection mechanism against this type of harassment has not been thoroughly investigated. As shown in Table 1, several platforms have implemented the personal space feature. Nevertheless, this feature is a passive defense against harassment. Since it restricts social interaction, users may elect to enable it only after harassment happens. Therefore, a more desirable mechanism should be able to detect potential harassment prior to its occurrence.

Real-World Challenges

In addition to the above technical challenges, we need to consider the following issues related to the physical world when designing the Metaverse. First, the Metaverse may cause ethical concerns. For example, it allows users to choose their avatars freely, but not all avatars are equally in demand. According to a study, users have a low demand for dark-skinned and female avatars (https://www.thenifty.com/race-and-nfts-636/; accessed on 25-Aug.-22), raising issues about race and gender representation in the Metaverse. Second, as the Metaverse becomes commonplace in our daily lives, user addiction will be a crucial issue (https://bit.ly/3RWEvzc; accessed on 25-Aug.-22). People may rely on the Metaverse to escape from the real world, as described in the novel *Snow Crash*. Beyond better regulation and guidance, how to effectively address this issue is still an open problem. Finally, virtual crimes in the Metaverse deserve our attention. The transactions in the Metaverse are conducted through blockchain-based NFTs and cryptocurrencies. Decentralization and non-regulation are the two main features of blockchain, which are prone to crime. In 2021, the worth of criminal activity regarding cryptocurrencies was up to $14 billion (https://reut.rs/3s0yEwC; accessed on 25-Aug.-22). Since the Metaverse could be a decentralized and free virtual world, our efforts to guide and monitor these issues in the real world may not be directly replicable in it. How to effectively address the above issues in the Metaverse deserves in-depth study.

Besides data privacy, harassment is another emerging concern in the Metaverse. The Metaverse has not only text-based or voice-based harassment, which has been studied for traditional social media, but also body movement-based harassment that can reflect the users' movement via their avatar through various sensors.

Conclusion

While the Metaverse has been deemed as a hypothetical NextG Internet, much of the discussion, in both industry and academia, has focused on its potential. In this article, after reviewing the current hype in the industry, we present the definitions of Metaverse, its enabling technologies, and our vision of its technical requirements. We then introduce existing social VR platforms that can be viewed as early prototypes of Metaverse. By measuring and comparing two representative social VR platforms, Workrooms and AltspaceVR, we point out the technical challenges and opportunities for future development. Given its multidisciplinary nature [3], we hope to see more initiatives emerging from not only the networking research community, but also other related disciplines such as social sciences, economics, computer graphics, AR/VR/MR, HCI, security, and privacy.

References

[1] B. Han, "Mobile Immersive Computing: Research Challenges and the Road Ahead," *IEEE Commun. Mag.*, vol. 57, no. 10, 2019, pp. 112–18.

[2] J. D. N. Dionisio, W. G. Burns III, and R. Gilbert, "3D Virtual worlds and the metaverse: Current status and future possibilities," *ACM Computing Surveys*, vol. 45, no. 3, 2013, pp. 34:1–38.

[3] S.-M. Park and Y.-G. Kim, "A Metaverse: Taxonomy, Components, Applications, and Open Challenges," *IEEE Access*, vol. 10, 2022, pp. 4209–51.

[4] B. Ryskeldiev, "Distributed Metaverse: Creating Decentralized Blockchainbased Model for Peer-to-peer Sharing of Virtual Spaces for Mixed Reality Applications," *Proc. Augmented Human Int'l. Conf. (AH)*, 2018.

[5] W. Zhang *et al.*, "DeepVista: 16K Panoramic Cinema on Your Mobile Device," *Proc. Int'l. Conf.World WideWeb (WWW)*, 2021.

[6] B. Han, Y. Liu, and F. Qian, "ViVo: Visibility-Aware Mobile Volumetric Video Streaming," *Proc. ACM Int'l. Conf. Mobile Computing and Networking (MobiCom)*, 2020.

[7] K. MacMillan *et al.*, "Measuring the Performance and Network Utilization of Popular Video Conferencing Applications," *Proc. ACM SIGCOMM Conf. Internet Measurement (IMC)*, 2021.

[8] P. Casey, I. Baggili, and A. Yarramreddy,"Immersive Virtual Reality Attacks and the Human Joystick," *IEEE Trans. Dependable and Secure Computing*, vol. 18, no. 2, 2019, pp. 550–62.

[9] F. Mathis *et al.*, "Fast and Secure Authentication in Virtual Reality Using Coordinated 3D Manipulation and Pointing," *ACM Trans. Computer-Human Interaction*, vol. 28, no. 1, 2021, pp. 1–44.

[10] L. Blackwell *et al.*, "Harassment in Social Virtual Reality: Challenges for Platform Governance," *Proc. ACM on Human-Computer Interaction*, vol. 3, 2019, pp. 1–25.

[11] R. Cheng *et al.*, "Reality Check of Metaverse: A First Look at Commercial Social Virtual Reality Platforms," *Proc. IEEE Workshop for Building the Foundations of the Metaverse (Metabuild), co-located with IEEE Confer. Virtual Reality and 3D User Interfaces (VR)*, 2022.

[12] S. Orts-Escolano *et al.*, "Holoportation: Virtual 3D Teleportation in Real-time," *Proc. Annual ACM Symposium on User Interface Software and Technology (UIST)*, 2016.

[13] R. Cheng *et al.*, "Are We Ready for Metaverse? A Measurement Study of Social Virtual Reality Platforms," *Proc. ACM SIGCOMM Conf. Internet Measurement (IMC)*, 2022.

[14] D. Koller *et al.*, "Protected Interactive 3D Graphics Via Remote Rendering," *ACM Trans. Graphics*, vol. 23, no. 3, 2004, pp. 695–703

[15] W. Zhang, B. Han, and P. Hui, "SEAR: Scaling Experiences in Multi-user Augmented Reality," *IEEE Trans. Visualization and Computer Graphics*, vol. 28, no. 5, 2022, pp. 1982–92.

Biographies

Ruizhi Cheng received a B.S. degree in Computer Science in 2021. He is currently working toward a Ph.D. degree in the Department of Computer Science, George Mason University, USA. His research interests include federated learning, networked systems, and virtual reality. He received the best paper award from IEEE MetaBuild 2022.

Nan Wu is a Ph.D. student at Department of Computer Science, George Mason University. His research interests are augmented/mixed reality, networked systems, immersive video streaming, volumetric video analytics, and security & privacy. He is recently focusing on privacy-preserving image-based pose estimation and network-friendly volumetric video streaming and rendering.

Songqing Chen is a professor of computer science at George Mason University. His research interests mainly focus on design, analysis, and implementation of algorithms and experimental systems in the distributed and networking environment, particularly in the areas of Internet content delivery, mobile and cloud computing, security, and distributed systems.

Bo Han is an Associate Professor of Computer Science at George Mason University. Before that, he was a Research Scientist at AT&T Labs. His research interests are in the areas of networked systems, mobile computing, and wireless networking, with a recent focus on immersive video streaming and XR.