

Optimal Latency in Collaborative Virtual Environment to Increase User Performance: A Survey

Shah Khalid
Department of Computer
Science, University Of
Malakand, Pakistan

Sehat Ullah
Department of Computer
Science, University Of
Malakand, Pakistan

Aftab Alam and Fakhrud
Din
Department of Computer
Science, University Of
Malakand, Pakistan

ABSTRACT

In this paper we evaluate Collaborative Virtual Environments (CVEs) in terms of network latency. Network Latency main requirements, given in the literature for CVEs, have been summarized. We also discuss effects of network latency on users' performance in the CVEs. Different factors can improve, by minimizing overall network latency, the performance of users in the CVEs. This review provides a comprehensive inside look to network latency in the CVEs and will help the researchers to adopt a network, with minimum latency, for their CVEs. The algorithms and ideas, suggested by the researchers, to be adopted by the CVEs to minimize network latency in virtual environment, are also discussed critically. Reducing network latency will ultimately improve the overall efficiency of CVEs.

Keywords

Collaborative Virtual Environment, Network Latency, Multimedia, Synchronization, Consistency, Throughput.

1. INTRODUCTION

Collaborative Virtual Environment (CVE) is new research field. Main aims of this study are to examine various aspects of CVE in the perspective of network latency to identify its various factors. The center of attention of this study is to be aware of the various research requirements for efficient CVE in terms of Network Latency. The ultimate goal of the study is to make improvements in the given techniques for minimizing the network latency or to suggest a new model or comprehensive algorithm for minimizing the network latency.

Collaboration has the ability to increase the efficiency of the group work and to divide the geographical obstacles between the peoples by using the virtual workspaces. Virtual workspace contain shared objects, its ranges from semi immersive to fully immersive virtual environments in which user can interact with each other like a real time fashion. Also the users can consider themselves to the part of the semi and fully immersive virtual environments [1]. For example, the strategy for a difficult job solution, like to design a building, to build a complex and large machine and to develop airplane engine, required experts to work in collaboration, well accomplished the said task easily. Similarly to contribute in a multi-player game and to make online tele-surgery required many peoples to work in collaboration. So to achieve a job in a real world, collaboration play very vital role. Researchers give stress on the development of the CVEs.

“Collaborative virtual environments are virtual reality systems that offer graphically realized, potentially infinite, digital landscapes. Within these landscapes, individuals can share information through interaction with each other and through individual and collaborative interaction with data

representation” [2]. “A CVE is a computer-based, distributed, virtual space or set of places. In such places, people can meet and interact with others, with agents or with virtual objects. CVEs might vary in their representational richness from 3D graphical spaces, 2.5D and 2D environments, to text based environments. Access to CVEs is by no means limited to desktop devices, but might well include mobile or wearable devices, public kiosks, etc.”[3].

By making this definition it is clear that although CVEs are usually associated with 3D graphical environments but this essential is not always be the case. “CVEs represent the computer as a malleable space, a space in which to build and utilize shared places for work and leisure. CVEs provide a terrain or digital landscape that can be ‘inhabited’ or ‘populated’ by individuals and data, encouraging a sense of shared space or place. Users, in the form of embodiments or avatars, are free to navigate through the space, encountering each other, artifacts and data objects and are free to communicate with each using verbal and non-verbal communication through visual and auditory channels”[3].

1.1 Requirements for CVEs

Some basic and general requirements for CVEs are the following.

Avatars: - The graphical representation for users in a CVE is called avatars. Different avatars are used based on objectives and task perform by the users. Humanoid, sphere, ball, simple virtual hand etc. are used as an avatar. The humanoid avatar uses the non-verbal channels for communication, like eye gazing, gestures, body orientation and movements etc. which enhance the sense of co-presences [4]. It can also increase the user performance, network traffic and the system complexity. While simple virtual hand reduced network traffic and system complexity but it has low level of social interaction in case of the complex task.

Audio/Video Support: - Mostly for teleconferencing applications voice telephony is used. When the latencies increase above 200 ms then it will lead to deterioration in the conversion. Video conferencing is face to face meeting and negotiation with each other's. In VR through avatars the sense of co-presence is created. Video based communication role is less important for collaboration in CVEs [5, 6, 7].

Characteristics of Data: - The transmission of data its storage and management in CVEs are affected by four attributes which are size of data, quality of service, persistence and queues [7]. In CVEs according to the objectives of the user performance these mentioned factors should be considered.

1.2 Types of Collaborative Work

In the field of CVEs to make collaborative interaction is one of the main challenges for researchers. The main objective of the research in field of CVE is to allow multiple users to make interaction with objects and to share a virtual space. In VR collaboration classification is the following [8, 9, 10].

Co-Presence CVE: - In this type of collaborative virtual environments the users recognize the co-presence through avatars. Also each user interacts with object individually. Whenever the user wants to make changes to the attributes of an objects or scene rendering, it will be visible to all his/her collaborators.

Active/Passive CVE: - In active/passive CVEs in which many users co-exist but only one user is active at a time and other users will be passive. The user which is active will take participation in interaction and manipulation process with objects in virtual environment.

Synchronous/Asynchronous CVE: - In synchronous CVEs via which two or more users can interact, handle and maneuver/manipulate the same object. This is also known cooperative operation. While in transmission process between users, whenever a user want to transfer the object from one location to another, then a second user moves it farther away, or when multiple users at the same time move an object is called asynchronous[11].

1.3 Applications of CVEs

The first CVE was developed for military training [12], and still a lot of work is doing in this area for example Institute for Creative Technologies at the University of California [13]. The CVEs has two main viable applications i.e. teleconferencing and multiplayer gaming. National Health Service (NHS) uses the teleconferencing facility for physician and patients meeting, training and various diagnoses purpose [14]. The downside is that when the distance is more and having slow network connection then latency issue is created. In result of high latency the awkward turn taking behaviors is created [15].

The delay issue is very critical in multi players distributed games. Due to this delay many forums and notice boards are dedicated to 'lag'. In a multi-player game like in 'EVE' and 'World of War craft' [16, 17] in spite of huge latency factor will support a large number of peoples, which restrict its effects by restrictive players to certain areas or by using some special mechanism to control the delay.

The only major obstacle which is still not solvable is latency. For example, while to develop the robotic supported surgery such as the Vinci Surgical System [18] in which the surgeon is 'remote' from the patient and still he/she is located in the same room. The training system for virtual tele-surgery has been demonstrated over a wide area network, but this relies on 'pseudo physics' to disguise the objects caused by latency [19]. True tele-surgery remains, for the moment, as unfilled goal.

To utilize the CVEs in field of engineering and medicine is still not clear yet. Some companies like Nokia now a day is in active to create devices which support haptic feedback in network games and mobile devices [20]. It means that advances in CVE technology will be used in entertainment industry in near future.

Up to now, haptic technology has been too expensive to accurately adapt into computer games, but this will change

with the introduction of the Novint Falcon [21] in 2007. We may also be bound to adopt the technology in the workplace too, rising oil prices and environmental concerns will make travelling long distances increasingly unviable, elevating the importance of CVEs for meetings and training exercises. As such, increasing our understanding of the behavioral impacts of latency remains a priority.

VR systems have developed as powerful tools to train people in tasks that are either very expensive or hazardous to copy in the real world [22]. The flight simulation is the well-known area where civil and military pilots are trained using virtual environments. Similarly, for assembly tasks training's and repairing task have already got the consideration of researchers, most of these tasks may be better achieved in collaboration [22].

In section 2 network delay is discussed. Section 3 contain thorough discussion of requirement for network latency followed by details of network latency in terms of consistency in section 4. We conclude the paper along with future work in section 5.

2. NETWORK DELAY in CVEs

CVEs allow group members at different places to work together in a natural way [23]. Delays occurs mainly from two sources, firstly, messages transmission via network and secondly to process the messages at the endpoints. Due to transmission and retransmission, switching, queue delay are arises. Due to type and network conditions the magnitude varies. Processing delays result from processing information at the sender, receiver and servers sides [24]. The members in CVEs are connected by means of a network which will create network delay when the members make interaction with each other's. In a closely coupled environment this delay causes inconsistency which considerably effects the actions and performance of collaborative users [25].

In CVE the members of the team will share information by making communication with each other's by sending messages. The actual information which is share and sends occurs in two forms i.e. discrete or continuous. For discrete information such as model updates or commands, only the order of messages is important. For real time or continuous information, time play very important role. In continuous stream of data, messages have temporal dependencies and placing the message in stream of data on his proper time has an effect on its interpretation. Streams are affected by two kinds of network delay. One is called latency and the other is jitter [26]. In CVEs delay can have severe effected due to collaboration/coordination feedback, communication, and understanding of the shared situation [24]. A lot of mechanism and approaches exist to deal with delay. To select appropriate distribution architecture and consistency mechanism delay will be reduced [27, 28].

There are two key statistics to be measured when characterizing the temporal performance of a network i.e. latency and jitter.

2.1 Latency

In CVEs when communication take place between team members then the interval between sending and the receiving of the packets among them is called latency. In network communication large latency occurs [29]. The main factors of latency in communication networks in CVEs are delay caused by the inconsistency among users, system architecture, network characteristics, congestion, node, queuing, processing and serialization delay, multimedia application, awareness and

time required for routing decision. These all factors jointly produce larger delay in communication. Latency is shown in fig. 1 (b).

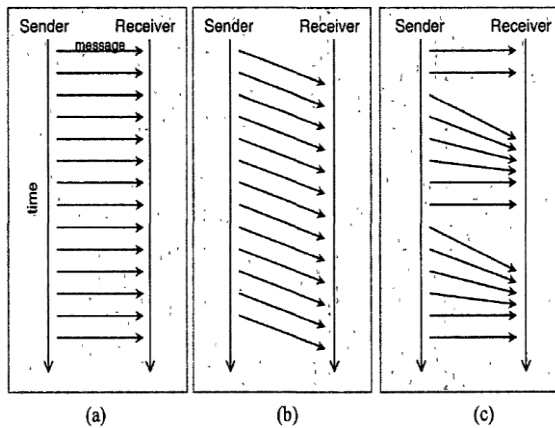


Fig. 1. Time-Series of messages from sender to receiver
 (a) No latency or jitter (b) latency but no jitter (c) jitter but no latency

At the network layer, CVEs use the internet protocol (IP), which is best effort protocol having no mechanism for flow and error control. At transport layer transmission control protocol (TCP) and User Datagram Protocol (UDP) are used. TCP is connection oriented and reliable protocol. TCP/IP can be combining used to provide reliable services to distributed group members in CVE environment. It ensures the integrity of data and also guarantees that the data has been received in order. However this reliability introduces additional latency. Checking for error detection and then apply some mechanism to correct the error will take some time. Also the protocol wait for acknowledgement of the packet, if it is received then it sends the next packet. Because of these extra processing TCP/IP is comparatively slow protocol. UDP protocol is connectionless and comparatively fast protocol than TCP. UDP does not guarantee the order delivery, error detection and correction technique. It sends the next packet without the acknowledgement of the first packet. UDP is fast as there is no extra processing involved. It potentially reduces the latency between sites [30].

2.2 Jitter

The change in latency from packet to packet is called jitter or simply variation in delay is called jitter and it causes spikes in latency [31]. It measures the difference between the arrival times of two or more messages. For example, if two messages are sent at 5ms interval, but the receiving interval varies from 5ms, this represents jitter. Jitter is shown in fig. 1 (c) in which different messages reach to its destination at different intervals. In CVEs system, distributed group members are connected through a network. The data is divided into packets and consecutive packets may follow different routes to destination. Therefore, different packets encounter different overheads and traffic conditions which causes jitter. One packet may be lost, which is to be resent. In case of in-order delivery, all the packet behind the lost packet has to wait till it is resent from the source. All these factors introduce jitter in the transmission. For distributed team members jitter appears as jerky moments. For example a tele-pointer will stick when data is delayed and will catch up moment when the data reach to destination machine. People can feel jitter up to 10ms. For audio and video applications, the multimedia data is first buffered and then playback begins which is smoother. It minimizes the jitter to zero level but increases the overall latency.

Jitter is not present in face to face communication sound waves take the same time in air to reach its destination. Jitter is characterized by two attributes. First, size of the delay and secondly the percentages of packets that are delayed. Jitter is a major issue in distributed groupware applications and can be minimized either at the network level i.e. by using quality of service parameters or application level by using buffering techniques. The second approach can be used for live communication as it does not provide real time service to distributed users.

In literature, small amount of work has been done to analyze the effects of jitter and latency on human performance in real time distributed applications. Vaghi et al. [33] analyzed the effects of delay on collaborative task performance in 3D virtual environments. According to Park [34], performance of collaborators is negatively affected by latencies of only 200ms. Another study conducted by Vaghi et al, [33] concluded that collaborative work become difficult at a latency of about 500ms.

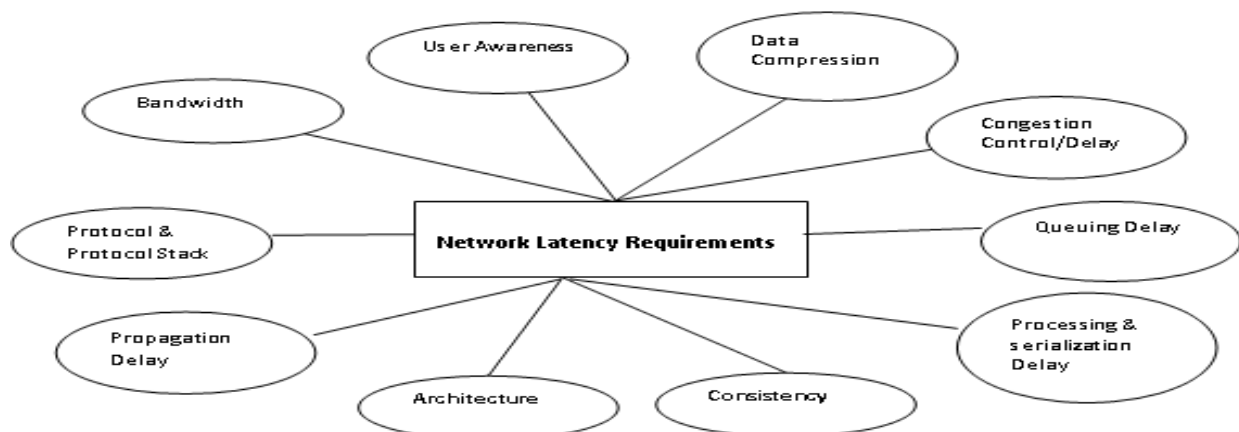


Fig. 2. Requirements for network latency in CVE.

3. REQUIREMENT FOR NETWORK LATENCY

Main requirements for network latency are the following as shown in Fig. 2.

3.1 Compression

Data compression is a mechanism to store and transmit data in transmission media efficiently [35]. The goal of the data compression is the same as that of relevance filtering, but uses a quite separate technique [36]. In CVE the main goal of the compression is that to minimize the size of packets which are transmitted in a communication media. The transmitted packets will be compressed at the sender side and then will make it decompressed in the receiving end. There are mainly two compression mechanism for data used [37].

Lossless Compression: - In this technique the size of the data will be minimize by changing the data encoding format and there will be no loss of data occur, for example Hoffman encoding schemes for data compression

Lossy Compression: - In this technique the redundant and irrelevant type of data will be removed whenever the packet are transmitted. Dead reckoning algorithm is a lossy compression algorithm. The storage and communication devices are greatly benefitted from data compression techniques. To minimize the message size the effective bandwidth of the communication channel is increased. We can send more data with low bandwidth when the data is compressed. When the size is reduced the space will be reduced for data transmission and the medium storage capacity will be increased. Also data compression makes storage devices faster as well. By using such type of compression algorithms which minimize the size of data will reduce the latency.

3.2 Congestion Control

In CVEs, whenever the subnet part become overloaded, then result the congestion ,because when router packets receiving is faster than forwarding of the packets by them, then one of two things can be result,

- The additional packets are stop by the router for entering to the congestion area until it will processed the stored packets.
- The congested routers will discard some queued packets to create space for new incoming packets.

Whenever the packets transmission via network reaches to the packets handling capacity of the same network, congestion occurs. It means that the packet arrival rate exceeds the outgoing link capacity. Insufficient memory to store arriving packets, slow processor, bursty traffic are concerned with efficiently using a network at high load are the main factors which causes congestion in network traffic [38]. For controlling the congestion many techniques are used. These includes warning bit, choke packets, load shedding, random early discard and traffic shaping.

Warning bit: - Whenever the router become congested then in the packet header a special bit is set to warn the source is called warning bit. The warning bit packet is send to the receiver. The receiver copy the warning bit and acknowledge the sender with warning bit. Whenever the source receive the packets with warning bits then the sender reduces its sending rate. The sender monitors the number of ACK packets it

receives with the warning bit set and adjusts its transmission rate accordingly [39].

Choke packets: - Whenever the router decides it is congested, it will send out choke packets. The choke packet tells the source of a packet arriving during warning state to slow down its sending rate [39].

Load shedding: - When the routers buffer memory becomes full then it will discard packets. The discarded packets chosen depends on the application and on the error strategy used in the data link layer. When the file is transferred, the old packets will not discarded because it will cause a gap in the received data. For real-time voice or video it is possibly better to throw away old data and keep new packets. Get the application to mark packets with discard priority.

Random early discard: - The router discards one or more packets before the buffer becomes completely full. The RED algorithm is used to compute the average queue length “average”, whenever the packets are arrived each time. There is a defined threshold value in RED algorithm, if the average is lower than this value then the congestion is called minimal or non-existent and the packets is queued. If the value is greater than the threshold upper value then the packet is discarded and it is called serious congestion. If the average is between the two thresholds then the congestion is onset and its probability is then calculated.

Traffic shaping: - This method is also used for congestion controlling. It “shape” the traffic before entering in the network. Mainly the rate in which the packets are send will be controlled in this method. Traffic shaping is used in ATM and Integrated Services networks. The traffic pattern is negotiated between the sender and the carrier at the time of connection set-up. Token Bucket and Leaky Bucket traffic shaping algorithms are used [38].

The first three techniques are used for congestion detection and the last two are used for congestion avoidance.

In CVE traffic flow become cumbersome then the data flow become small and will take more time for data transmission. Ultimately the latency will be increased when the congestion is high. So to make the latency low then the congestion will be avoided or to make it minimum. So to improve the user performance in CVE the latency level is kept low by controlling the congestion.

3.3 Users Awareness

“It is the knowledge of a user about intensions, action, feelings and status of the other users in CVEs in the shared space”. It measures the extent, nature or quality of interaction among the objects or users [4]. It has mainly three types.

Social awareness: - To maintains information about others in a social or informal context is called social awareness. For example, knowing about a person’s interest and intention. Through gaze and facial expression it is achieved.

Informal awareness: - In the work community whenever a person generally sensed the peoples is called informal awareness. It facilitates limited and casual interaction.

Group structural awareness: - Knowledge about people’s status, roles and responsibilities.

Workspace awareness: - Knowledge about people’s location, identity and activities is called workspace awareness.

In CVE when the awareness level among the users is high then the latency will be low and ultimately the user performance will be increased.

3.4 Propagation Delay

Propagation Delay can be defined as “the amount of time it takes for a packet dispatched by the sender to reach the receiver’s computer at the application layer level” [41].

In a data communications system, propagation delay refers to the time interval between the leaving of a signal from the source and the arrival of the signal at the destination. It range is from a few nanoseconds or microseconds in local area networks (LANs) up to about 0.25 ns in geostationary-satellite communications systems. Extra propagation delays can occur due to the time required for packets to make their way through guided media and nodes of the Internet [42].

Table 1. Propagation Delay and Distance [40]

Distance	Propagation Delay (milliseconds)
1 mile	8.2 microseconds
5 miles	41 microseconds
20 miles	0.164 ms
100 miles	0.82 ms
200 miles	1.64 ms

Propagation delay depends upon time as shown in table 1, when the network devices like router, cables etc. take more time for propagation of packets and hence ultimately the latency will be increases.

3.5 Processing and Serialization Delay

Processing delay refers to the time taken to analyze the packet and accomplish the transmission from source node to the destination, through the intermediate route. This includes application processing and operations on the relaying network nodes. In communication system, the network devices receive packets, then processed it and finally forwarded. Doing this process finite amount of delay will occur because of the value added features of networking devices and add additional delay. Those features that support the hardware assistance will reduce latency. For example LAN Switch and WAN router have been converging. Data communication industry make transmission from TDMA bases serial lines to High-speed metro Ethernet (uses hardware assistance).

Serialization is the conversion of bytes (8 bits) of data stored in a computer’s memory into a serial bit stream to be transmitted over the communications media. Serialization takes a finite amount of time and is calculated as follows,

Serialization delay = packet size in bits/transmission rate in bits per second.

For example, Serialization of a 1500 byte packet used on a 56K modem link will take 214 milliseconds. Serialization of the same 1500 byte packet on a 100 Mbps LAN will take 120 microseconds. The links which operate on lower transmission rates, serialization can represent a significant delay. Mostly links this delay is a small fraction of the total latency when compared to the other contributors. In audio/video data streams generally use tiny packet sizes (~20 ms of data) to minimize the influence of serialization delay [42].

Serialization delay can be minimizing by using the cut through switch because the time will be saved. Cut through switch is the switch in which “the start transmission of packets out on the destination port before it has received the full packets on the incoming port”. With cut-through switching, save the time it takes to transmit/receive the entire packet. In early days of transmission of the data like 10 Mbps LAN links this meant a lot. The time it takes to transmit a packet at 10 Mbps is between 51.2 and 1200 microseconds for a 64 or 1500 byte packet. Today, at 1 Gbps, this drops to between 0.512 and 12 microseconds. At 10 Gbps, it further reduces to between 0.0512 and 1.2 microseconds. Table 2 summarizes serialization delay effects associated with various link types for 64 and 1500 byte packet sizes [40].

Table 2. Serialization Related Summary.

Packet Size	Link Size	Serialization Delay
64 bytes	256 Kbps	2 ms
	1.5 Mbps	0.35 ms
	100 Mbps	5.1 microseconds
	1 Gbps	0.51 microseconds
	10 Gbps	0.051 microseconds
1500 bytes	256 Kbps	46.98 ms
	1.5 Mbps	8 ms
	100 Mbps	120 microseconds
	1 Gbps	12 microseconds
	10 Gbps	1.2 microseconds

When the networking devices did not uses the value added features, the latency will be reduced.

3.6 Queuing Delay

When the packets are assigned to a queue for transmission, it stays in the queue for some time. This time is called queuing delay [43]. In the packet switched networks the queuing delay has the most adverse effect on packet delay. Any packet moving through the switches in its path from source node to destination node without experiencing any queuing delay has the minimum end-to-end delay [44]. Queuing delay occurs at the output ports as the data of many input ports may be destined for the same output port [45]. There will be no queuing delay if the speed of input link is equal to the speed of output link and only one input link feeds an output link.

Message buffering occurs because of burst traffic arrival i.e. when the output port cannot forward all the input data continually so it is stored in the queue. Depending on knowledge of the input traffic patterns we can analyze the buffering delay [44].

Queue delay has a great effect on the performance of switch. If we minimize this delay the overall performance of the switched networks will be increased and the overall end to end delay will be minimized. Network congestion can be controlled if we minimize the packets queuing time. In other word we can minimize the congestion if we make our network in such a way that can handle traffic burst. The preferred strategy is to avoid congestion in the first place [40]. Quality of service methods such as Low Latency Queuing (LLQ) should be used when some level of congestion is expected in the network. Overall QoS strategy is required when different types of traffic are moving through the converged network. Weighted Random Early Detection (WRED) method may not work in this case as it randomly drops packets when we expect congestion by looking at the Differentiated Services

Code Point (DSCP) bits. In response to this situation the Transmission Control Protocol (TCP) will detect the drops and adjust its sliding window size. This will increase the overall latency of the communication system. However, it might be better than experiencing packet loss and retransmissions.

To avoid the queuing delay the congestion should be avoided by doing this the latency will be reduced.

3.7 Bandwidth

In communication, the data bandwidth is same to the transmission rate being used. For communication point of view the bandwidth is very important because it is the maximum size/capacity of a data link. It's very important to know that the data bandwidth is "it is the transfer of given data segment over a given period of time to obtain maximum data flow is called bandwidth" [42].

As for as the latency is concerned the data bandwidth doesn't directly effect on it, but the number of hops between the sender and server to which the sender is connected can affect the ping time dramatically. Each hop in the form of router can create a routing delay, so usually we want a such a network route which contain a small number of hops (it is not always the case, since there is speed difference between the networks, because some networks may be faster than others) [46]. So in data transmission of small packets with minimum numbers of hops, the transmission speed will be high and ultimately the latency will be low. However in small packet transmission extra overhead is created due to the packets parameter inclusion for every packets, like IP header, Ethernet header etc.

Optical fiber is used for wideband data transmission. Optical fiber media offers a large bandwidth up to tera hertz. This media can carry many tera bytes per second of data speed and capacity. So for huge bandwidth data transmission optical fiber media is the best suitable choice. Hundreds of video can be transfer simultaneously in this media. It is just like highways which can carry different types of data and communication links pass to it with tremendous speed. Using this huge capacity media for data, the latency will be minimum.

However the guided media like optical fiber can be used only for end to end communication but it is not appropriate for broadcasting purposes. So for broadcasting purposes wireless media has an advantage over optical fiber. Wireless media is used to connect too many peoples in remote areas. To connect too many peoples from remote areas are very costly and also a lot of geographical terrain problems involved for which the guided media is not suitable. So the best choice for such a circumstances is the unguided media. Similarly for traveling users unguided media is used.

Wireless communication includes many forms ranging from simple amplitude modulation (AM), frequency modulation (FM) to the highly technical new techniques like satellites communications. For remote areas wireless communication are the boon because it is possible to provide services to the fewer users at a much lower rate. But the data caring capacity of wireless is slow as compare to wire media [47]. Due to the low data caring capacity of wireless media the latency will be high.

3.8 Architectures of the CVEs

Architecture/Model is something in abstract form which presents something to viewers. So Data model for

collaborative virtual environments may be the most difficult choice to build a CVE is to define where to place the data related to objects and state of the virtual world. This decision will affect the communication requirements for CVEs, The reliability of the data and scale for CVEs [48]. The important models for CVEs which are described in [7, 48] are shown in fig. 3.

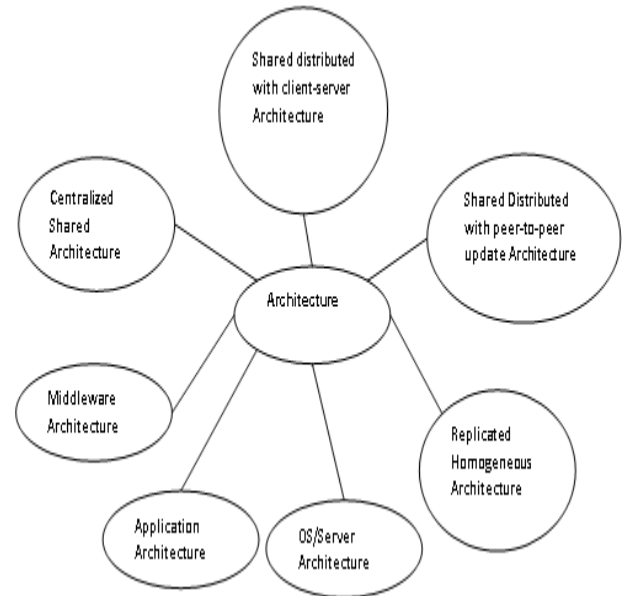


Fig.3. Different Architecture

3.8.1 Middleware Architecture

Features of messaging protocol is another factor that causes latency in the networking environment. The efficiency of the protocol need to be optimized during setup, retransmission and tear down for better performance.

The loquaciousness of the protocol affects the number of packets on the lines and also causes some latency issues. The middleware determines the number of multi cast group that an application will use. Dividing the data into logical groups by the developer minimizes, up to some extent the latency and improves efficiency [40].

3.8.2 Application Architecture

It is very challenging in the design of application architecture to determine how to extend the trading application without increasing the overall system's latency. The application's number of tires, distributed vs. centralized processing, and event-driven vs. service-oriented architecture are some typical trade-offs that need to be considered. The trade-off between efficient execution and rapid development cycle must also be made by the application developers.

Another common mistake that application developers make is testing code usually on a single subnet in lab and do not communication with the networking section till the application is near to go into production. This lack of coordination between the two teams may results inefficiencies in the performance of application or the network design. The data store's type and location is also an important factor.

Grid computing in application architecture is a usual solution for handling of the above discussed problems. It reduces application's processing time by the parallel execution of different parts of the application or by deploying multiple physical servers. The communication gap between the application developers and networking department may cause

network latency. In order to minimize network latency, grid computing comes into action by reducing the processing time of applications via real time memory data base products [40].

3.8.3 Operating System Architecture

Network latency can also be affected by software components such as operating system (OS) and hardware components such as the CPU, memory, and hard disk. The access times for different RAM technologies varies from 9 to 70 nanoseconds. The CPU must copy data between application buffers and network buffers in a conventional network stack implementation. This overhead is combined with the fact that memory is much slower than CPU in speed. For instance, the speed of processors such as Intel Xeon is reaching 4GHz whereas speed of RAM chips is only 400MHz with 10:1 ratio in clock speed. This wastes the precious time of processor in waiting for RAM for 10 clock cycles to fetch and send the data to the processor when the data is not available in its cache memory.

Also static RAM and Dynamic RAM are the two types of RAM technologies. Static RAM is constructed with transistors and is better in speed than Dynamic RAM (based on capacitors) which is suffered from periodic discharge. Advanced hardware technologies such as multiple CPUs (dual and quadcore), 64 bit processors, fastest memories, fastest disks, etc. can be helpful in reducing system's overall latency [55]. Others architectures like client server, peer to peer, homogeneous replicated etc. are discussed under consistency.

4. CONSISTENCY IN CVEs

Collaborative virtual environments (CVE) are designed to allow people in remote locations to work together over networks. In CVEs the people can work together on designing systems, complex group task performing, learning from each other's and share collaborative experiences through these shared virtual environments. To share the same information in all sites, all CVEs transmit information about their local entities to remote sites through network [49, 50]. To keep this consistency, all the users will be renders in real-time, so that the user will not notice any difference between local and remote entities in the environment. Thus, to maintain the real and natural interaction among the users on the network, the CVEs required a high quality of services (QoS). For example, in CVE the users expect a precise visual scene of the remote object's movements to avoid collisions between their objects and those controlled by the remote partners. To ensure the global consistency, efficient collaboration between the users are required. In CVEs the users are dispersed in different physical location. The CVEs system have to guarantee the VE consistency, despite the low bandwidth and network latency issue.

The consistency of CVEs can be characterized by synchronization, causality and concurrency mechanisms as shown in fig. 4. Synchronization is divided into time and spatial synchronization.

When order of all users events in CVEs are same for all users, is called causality. While Concurrency should be maintained. Conflicts will be produced when the same arguments of the virtual objects will be changed at the same time. In CVEs the users own modification will be avoided which produced inconsistent state of the virtual objects.

Consistency and system responsiveness are directly linked with each other. Responsiveness means the time required by the system to respond to user actions. During the interaction process of the data responsiveness can be quantified by the

system latency, i.e. the time among a user action and the system response. This type of latency is created due to the processing delay of the events and the transmission time over the network. To improve the consistency of a CVE will ultimately increase latency during interactions and vice versa. So, CVE systems must reach a tradeoff between system responsiveness and consistency [51].

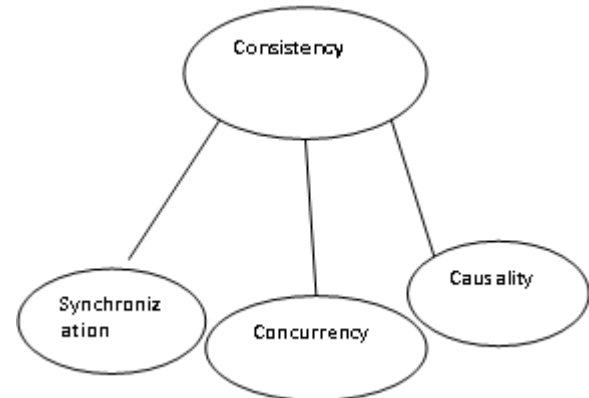


Fig. 4. Consistency parameters

The main consistency maintenance mechanisms are the following.

4.1 System Architecture

The CVE performance and its consistency are associated with its system architecture. For example, there are some architecture exist which maintain strong consistency but during interaction there latency is high. In contrast, other architectures accept a few inconsistencies but offer a better responsiveness during interactions.

4.1.1 Network Architecture

All CVEs system is connected by means of broadcast, multicast and unicast data transmission method and interaction will take place between the nodes of the CVEs. Peer-to-peer, client/server and hybrid architectures are used in CVEs [52].

Peer-to-peer architecture: - In peer-to-peer architecture events are transmitted directly from one node to another node so fast communication will occurs between the nodes. So, it enables a few users to have a closely coupled interactions and strong synchronization. However, increasing the number of users will increase the number of messages transmitted on the network. To transmit all the changes in VEs, It is difficult to contact all the nodes at the same time. That's why time synchronization and global consistency of the CVE may be difficult to ensure.

Client/Server architecture: -In this type of architecture the clients are connected to a centralized server. The server stores and manage all communication among the nodes of the network. In this architecture the server contact all the nodes at the same time. So, time synchronization and CVE consistency are easier to maintain than with the peer-to-peer architecture. However, when two users want interaction simultaneously with each other than latency will be increase. Bottleneck occur when the user increases because the number of request increases on server side and ultimately all the communication process will become slow down.

Hybrid Architecture: -Due to latency and slow down communication problem in server architecture and time synchronization and global consistency maintenance in case

of many users in peer to peer architecture the hybrid network architecture is used. This architecture uses both peer-to-peer connections and one or several servers. For example, to make the communication fast, using peer-to-peer connections and for maintenance of better consistency using server.

Another hybrid architecture exist to facilitate collaboration between nearby users (according to their location in the VE) by reducing the latency between them. All stations are connected via a central server, when the user come closer to each other in VE then a peer to peer connection is established which will increase the VE consistency [53].

4.1.2 Data Distribution

Shared centralized world: - In central server all the relevant data are stored about the virtual environment. Similarly, on this central server all CVE objects behaviors are executed. To modify an object by the user it will send request to the central server, server process its request and then send the up-to-date state of the object information to all participating nodes in the CVE, including the one that asked for this modification. This method keeps consistency among all the nodes and avoids replication of data. This method has two main drawbacks. Firstly Latency can be increase when transmission delays occur between the clients and the server during the interaction process. Secondly with many users, a bottleneck can appear on the server because it has to send updates to all the nodes at the same time especially with unicast connections.

Homogeneous replicated world: - In this strategy all nodes of the system in the same data warehouse will be initialized, which contains all the information such as terrain, geometric models, textures, object behaviors, etc. about the virtual environment. In most case the data will be available locally to the user when he log in to the system Otherwise, he can obtain the data from a server. During every session, the database of the system evolves independently on each node which additionally requires a synchronization mechanism to control the executions of object behaviors on each node. For the purpose of consistency, object medications and some special events such as collision between two objects are transmitted on the network which enables all nodes to update their database.

This data distribution policy has two main advantages. Firstly, because only update messages are sent, the number and the size of messages transmitted on the network are really small. Secondly, virtual object modifications are performed locally. Therefore, latency is very low during user interactions. Data replication may introduce inconsistencies between users of VEs because of delays or loss of data when updates messages are sent over the network. Also additional mechanisms must be provided to manage the concurrent access of objects on each node. The user is able to perform a local modification of an object, but modification conflicts are only checked when the changes are transmitted to the other nodes of the system [54].

Partially replicated world: - This is the hybrid approach which combine the concepts of totally centralized and copied data distributions. In Some CVE systems chooses this hybrid solution in order to avoid the drawbacks of these two approaches and to utilize their advantages.

4.1.3 Communication protocols

The protocol is used to ensure integrity and consistency of data while transmitting in collaborative virtual environment.

Protocols are basically a set of technical rules for the transmission and receipt of information between computers in network. Most of the protocols are implemented on application, transport and network layer of the OSI model. To carry out CVEs, protocols are usually implemented either on application layer or transport layer. IP is used to implement the logical link between the devices and is used in network layer of the OSI model.

Classical Protocols: -Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) are use on transport layer of the OSI model. They are also known as classical protocols. TCP is used only in one-to-one broadcasting. UDP on the other hand is used for multicasting and broadcasting. TCP is slow but it is likely to be more accurate because it uses an acknowledgement from the receiver to sender. In CVE a persistent interaction is used, which need a huge amount of data or regular information exchange, it will use TCP. While using TCP the reliability of information is assured but on the other hand it introduces additional latency because of its acknowledgment feature and also because of its large amount of data carrying responsibility which creates congestion on transmission line [55].

UDP is used for fast data communication. It sends data without any acknowledgment from the receiver. UDP send the data in non-connected mode so to calculate correct ordering of packets after receipt and error correction is difficult. UDP is used in multi-cast and broad-cast data communication. When the small amount of data is sent and received in an arranged time then it is called punctual interaction. For CVEs having punctual interaction, UDP is used [56]. The message is resent when it is lost to keep the state of the object up-to date in CVE like SPLINE [57] and SIMNET [58]. Reliability in UDP is decreased but it offers low latency.

Multicast Oriented Protocols: - In the beginning, the CVE system uses either unicast or broadcast system for data transmission. For example the MRTToolkit [59] uses the unicast transmission method for data sending while the SIMNET [58] were uses the broadcast system for transmission. In an environment when the users are large enough then unicast is not the best solution for data sending. Broadcast is also uncertain when a lot of nodes connected to a network. To resolve these problems a new technique is used called multicast data transmission. It sends data to many users at the same time [56], this solution has several disadvantages.

For point-to-point communication efficient multicast implementation is difficult. Many routers do not support multicast. The group in multicast can be limited because of the addressing, administration and congestion control.

Virtual Reality Dedicated Protocols: - Real Time Protocol (RTP) which is used for audio and video data transmission and for simulation of data in virtual reality environment [60]. RTP supports both unicast and multicast transmissions. A variation of RTP for interactive application called RTP/I. RTP\I uses different flavors of packets to manage event communications, object states transmissions, state changes, and state queries [61]. To support virtual reality modeling language (VRML) Virtual Reality Transfer Protocol (VRTP) is used, in the same way that HTTP is the HTML support [62]. It is the extension of the HTTP protocol. The main goal of the VRTP is to manage 3D interactive objects and hence to meet the requirements of CVE. In VRTP a node can take the role of a server, client and/or a peer.

To exploit Internet for virtual environment a specific application protocol is designed known as Distributed Worlds Transfer Protocol (DWTP) [63]. It is based on standard protocols such as TCP/IP and UDP/IP. In DWTP Events are used for consistency, messages are used for joining or leaving the virtual environment, files are used for the transportation of 3D scene or object geometry, and data streaming is used for continuous audio and video transmission.

The DWTP concept is based on daemons and participants. Daemons provide services to the participants, detect transmission failures, transferring lost data, transmit the virtual environment content to new participants, and extend the architecture for participants unable to use a multicast channel.

Industrial Environment Specific Protocols: - Classical and VR dedicated protocols are not used for some industrial applications due to the modern security constraints. For examples the ShareX3D firewall is used for which HTTP is required [63]. So ShareX3D uses the client/Server Architecture. As HTTP does not allow requests from the server to the client, therefore, to resolve this problem a technique is used called “long polling” technique. Long polling technique introduces latency because it requires reestablishing the connection frequently between the client and server.

Many options are used in protocol stack which can affect the overall efficiency of the data delivery. For that we need to know the characteristics of the version which are used by the devices and check its compatibility with other stack option and version, if compatible then latency will be reduced. For example the algorithm of Nagle’s is used for minimizing the overhead of the network by using the method of packets concatenation, but this algorithm interacts very badly with TCP delayed acknowledgement. However checking its compatibility with others stack and version will create additional overhead [40].

4.2 Synchronization/Time Management Techniques

Time is an important element of a CVE. The concept of time can differ from application to application [65].

Time and consistency have very important relationship. In a perfect consistent CVE, all the users perceive the same state at the same absolute time. However, due to the network latency this perfect case can never happen [65]. Different solution exist to improve consistency over time according to the required responsiveness for interactions.

4.2.1 Lockstep Synchronization

The lockstep synchronization used in RING or Open MASK is the easiest way to ensure the consistency of a CVE [67, 68]. “It stops some of the participating nodes go to the next step until all nodes have processed the current simulation step”. No node is allowed to increment its logical clock until all participating nodes have acknowledged that they are ready now for next simulation step. So all events will be processed in correct order and no need of roll back is required. It guarantees consistency but not in real-time. If there are delays or losses during transmissions, the time spent to wait increases, and the system responsiveness breaks down. Furthermore, simulation steps are not necessarily constant, so the system jitter can be substantial.

4.2.2 Time Warp Synchronization

Jefferson proposed the “Time Wrap” synchronization technique, which is an optimist technique, in which each event is processed in the time of arrival. All the events are marked with a “timestamp” [66]. Whenever an event of older “timestamp” is received, then the event which is just been processed, the time wrap synchronization cancel the current ongoing events processing with a most recent timestamp. To catch up the current time will processed all these events. Furthermore it sends messages to cancel the incorrect messages which are send during the roll back propagation process. By using this synchronization method the latency interactions will become low. This method is used in that situation when the roll-back happens rarely, because they are extremely frustrating for users. Several systems propose to quickly display several key-frames during the re-execution of events to facilitate the users’ understanding. Finally, this method needs to store the received events to re-execute them in case of roll-back [66].

4.2.3 Predictive Time Management

This method will predict the events in advance before their occurrence and send them on the network. This method is used in PARADE system to manage consistency in CVE with inherent latency [69]. As in all virtual environments some objects are not predictable so this mechanism can’t be applied everywhere. Particularly, in that situation in virtual environment to predict user actions. To detect collision this method is used by PARADE. Locally all the events are predicted, marked with a “timestamp” and sent to other nodes, where these events will be processed at the appropriate time. This predictive management is interesting only if the time between the sending of the predicted event and its processing is higher than the network latency. Otherwise the message will arrive too late to be processed.

4.2.4 Server Synchronization

In server/client architectures, the server using logical clock to synchronize events for example in ShareX3D [68], for each object of the CVE the server keep a “state number”. Whenever the server receives a message for change the object, it increases the “state number” of the objects.

4.3 Concurrency Control

The centralized data distribution mechanism in which server is used to control overall communication between users of the CVE give us the guarantees of an implicit control of concurrent access to CVE objects. Similarly, is the system of referent/proxies paradigm as used like in Open Mask and which only referent of an object can be modified by users [68]. However in replicated homogeneous mechanism the users can access and modify objects locally before these changes are transmitted to other users. So mechanisms are required for user management explicitly, to avoid the inconsistencies in the virtual world because of the concurrent access to the objects by users. Three mechanism are used to manage the concurrent access of the users [50].

Pessimistic mode: - In this mode of interaction with object in VE, only one user can modify an object with the same time with a lock mechanism like in BrickNet [71]. In VE whenever a user want to perform the manipulation techniques, so he/she will asks to become the owner of the object. An object will have only one owner, if the object is already with other owner, then the current user will wait until he/she release the ownership of the object. So in this mode only the current owner will manipulate the object. So by using

this method no concurrent access of an object can occur. Whenever the number of users and network latency are high. So to acquire the ownership of the object will required a lot of time and hence the latency will be introduces during the interactions process. BrickNet uses server for this mode of interaction to save the information that which one is the owner of each object [71]. This mode is not suitable in peer-to-peer architecture because when a node requests the ownership of an object, it must ask all the other nodes if they own this object.

Optimistic mode: - This mode permits users to alter objects without checking the possible concurrent access on these objects. In this method low latency interactions with object will be produced. However, when a conflict occurs, it is necessary to make a correction. This a complex task to resolve the conflict and will require that the users perform their action again. So to perform the action again and again the extra latency overhead is created.

4.4 Information Management Techniques

Predictive Contract Agreement Mechanisms Dead Reckoning: - “Predictive contract agreement mechanisms are optimistic consistency maintenance mechanisms that operate by employing a form of controlled inconsistency” [72]. In this approach participants agree on a prediction algorithm, a related threshold error and a convergence algorithm. Threshold error reflects the amount of inconsistency that will be allowed to occur between the true state and the predicted state. When threshold error is exceeded, the inconsistent state is corrected by transmitting current state information to all other participants who must subsequently converge to this updated state. Predictive contract agreement mechanisms reduce the amount of network traffic and therefore reduce the network latency.

For bandwidth saving mechanism dead reckoning is used. Dead reckoning mechanism also compensates for variable communication latency [73, 74].

Dead reckoning has been used to predict the occurrence of deterministic events. The deterministic events can then be transmitted to other users in advance and thus improve the consistency [75, 76]. Pre-reckoning algorithm which complements the dead reckoning approach is also used [77].

Relevance Filtering: - Mostly users of the CVEs are only interested in the subset of large volume of the available data. Relevance Filtering is also based on this idea. The main purpose of the relevance filtering is to control the flow of data. There are many model used for data flow in CVE but the most general is the aura-nimbus model [78, 79]. The main components of this model are the aura, the medium, the focus and the nimbus [80]. The aura is an object- and medium-specific sub-space in which interaction (communication) may occur. Medium is used for communication like in the form of audio, visual or text. Various architectures are used for relevance filtering implementation. For example in RING system, Server client architecture is used. The server is used to forward the updates to those nodes with entities which are visible to each other [81]. Similar to this approach is Area of Interest (AOI) data management [82]. AOI technique provide customized stream of data for each node. Another architecture uses the unicast data transmission mechanism. In unicasting when the user leave the group then communication is stopped. Multicast communication architecture is another technique which are used for data transmission. In this case data is transmitted to multicast groups and any time new nodes can

join the group are leave the group to start or stop receiving data from that multicast group. In addition the number of join/leave operations must be minimized as these increase network traffic, consume computing resources and hence cause latency to increase.

Packet Bundling: - For data management, Packet Bundling or aggregation technique is used. In this technique the network can only handle a small limited number of packets per unit time [83]. In packet bundling the number of packets transmitted are reduced, in contrast to packet compression, in which the packets size will be reduced during transmission. Packets bundling assembled a number of individual packets into a larger data unit and transmitting this new unit as a single packet. By this method the consistency will be increases but latency will also be increases.

In CVE when the data and the users are consistent then the user performance will be efficient and latency will be reduced. If the information and users were inconsistent then the latency will be increased.

5. CONCLUSION AND FUTURE WORK

Collaborative Virtual Environment enables people to work combine through the use of different technologies. In CVE systems the members work together which increase the overall efficiency of the work for a given organization. The network infrastructure is the main requirements for CVE system to make it efficient. In data transmission in networking environment the issue of latency is occur which will affect the performance of the users in CVE. The main causes of latency are the congestion delay, queuing delay, propagation delay, processing and serialization delay, inconsistency between the information and users, security of data, data compression, storage equipment's, multimedia affects and users awareness. To minimize the latency these issue will be resolved as explain in literature. When the users in CVEs are aware, they are synchronized, uses simple and powerful processing algorithm for compression and security of data, Use fast storage devices in networking, avoid the congestion, node, queue, processing and serialization delay, which will ultimately minimize the latency. By minimizing the latency the overall performance of the users in CVEs will be improved. In the above main factors for latency further research are required to provide more efficient consistency maintenance, congestion avoidance, secure, compression and decompression algorithms to minimize the latency and will improve the user's performance.

6. REFERENCES

- [1] A. Alam, S. Ullah, S. Khalid, Fakh, I. Rabi, “Computer Supported Collaborative Work (CSCW) and Network Issues: A Survey”, Information Journal Tokyo Japan, pp. 7995-8020, 2014.
- [2] Gibson, W. Neoromancer. Grafton Books, 1989. Gaston K J. Global pattern in biodiversity. Nature 2000; 405(1):220-7.
- [3] Churchill, E.F. and Snowdon, D. “Collaborative Virtual Environments: An Introductory Review of Issues and Systems. In Virtual Reality: Research, Development and Applications”, 3, pp3-15, 1998.
- [4] Roberts, D., Wolff, R., Otto, O., Kranzlmüller, D., Anthes, C., and Steed, A. “Supporting social human communication between distributed walk-in displays”. In VRST '04: Proceedings of the ACM symposium on

- Virtual reality software and technology, pages 81–88, New York, NY, USA. ACM 2004.
- [5] Greenhalgh, C. and Benford, S. Massive: “A collaborative virtual environment for teleconferencing”. *ACM Trans. Comput.-Hum. Interact*, 2(3):239–261, 1995.
- [6] Frécon, E. and N’ou, A. A. “Building distributed virtual environments to support collaborative work”. In *VRST ’98: Proceedings of the ACM symposium on Virtual reality software and technology*, pages 105–113, New York, NY, USA. ACM, 1998.
- [7] Leigh, J., Johnson, A. E., and DeFanti, T. A. “Issues in the design of a flexible distributed architecture for supporting persistence and interoperability in collaborative virtual environments”. In *Supercomputing ’97, Proceedings of the 1997 ACM/IEEE conference on Supercomputing (CDROM)*, pages 1–14, New York, NY, USA. ACM, 1997.
- [8] Margery, D., Arnaldi, B., and Plouzeau, N. “A general framework for cooperative manipulation in virtual environments”. *Virtual Environments’99 Proceedings of the Eurographics Workshop*, pages 169–178, 1999.
- [9] Otto, O., Dave, R., and Robin, W. “A review on effective closely-coupled collaboration using immersive CVEs”. In *VRCA ’06: Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications*, pages 145–154. ACM, 2006.
- [10] Aguerreche, L., Duval, T., and Arnaldi, B. “A description of a dialog to enable interaction between interaction tools and 3D objects in collaborative virtual environments”. In *Proceedings of VRIC 2009*, pages 63–73, 2009.
- [11] Pinho, M. S., Bowman, D. A., and Freitas, C. M. D. S. “Cooperative object manipulation in collaborative virtual environments”. *Journal of the Brazilian Computer Society*, 14:54 – 67, 2008.
- [12] S. Singhal and M. Zyda. “Networked Virtual Environments Design and Implementation”. *ACM Press SIGGRAPH Series*, Addison Wesley, 1999.
- [13] S. Noll. “Entertainment technology and virtual environments for military training and education”. In <http://www.educause.edu/ir/library/pdf/ffp0107s.pdf>.
- [14] NHS Medical Illustration Teleconferencing. <http://www.ich.ucl.ac.uk/ich/html/support/medill/teleconferencing>.
- [15] K. Ruhleder and B. Jordan. “Meaning- making across remote sites: How delays in transmission affect interaction”. In *Sixth European Conference on Computer Supported Cooperative Work ECSCW*, pages 411–429, Copenhagen, Denmark, 1999.
- [16] World of Warcraft. <http://www.worldofwarcraft.com>.
- [17] EVE Online: “The World’s Largest Game Universe”. <http://www.eveonline.com>.
- [18] “Intuitive Surgery da Vinci Surgical System”. <http://www.intuitivesurgical.com>.
- [19] C. Gunn, M. Hutchins, D. Stevenson, and M. Adcock. “Using collaborative haptic in remote surgical training”. In *world HAPTICS First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 481–482, Pisa, Italy, IEEE Computer Society, March 2005.
- [20] “An Innovative Tactile Feedback System for a Personal Mobile Telecommunications Device”. <https://tkp.tkk.fi/projects/2004/nokia>.
- [21] The Novint Falcon. <http://www.novint.com/falcon.htm>.
- [22] Richard J. Adams, Daniel Klowden, B. H. “Virtual training for a manual assembly task. In *Haptic-e*”, volume 2, 2001.
- [23] Tatar, D., Foster, G., and Bobrow, D, “Design for Conversation: Lessons from Cognoter. *International Journal of Man-Machine Studies*”, 1991. 34(2).
- [24] Carl Gutwin et al, “Revealing Delay in Collaborative Environments”, Vienna, Austria, April 24–29, 2004.
- [25] Salvador, T., J. Scholtz, and J. Larson, “The Denver Model for Groupware Design. *SIGCHI Bulletin* 28(1)”, 1996.
- [26] Fluckiger, F. “Understanding Networked Multimedia”, Prentice Hall, 1995.
- [27] Phillips, W.G. “Architectures for Synchronous Groupware”, Technical report 1999-425. Department of Computing and Information Science, Queen’s University, Kingston, Ontario, Canada, 1999.
- [28] Greenberg, S., and Marwood, D. “Real Time Groupware as a Distributed System: Concurrency Control and its Effect on the Interface”, *Proc. CSCW’94*, 207-217, ACM, 1994.
- [29] A. Alam, S. Ullah, S. Khalid, Fakhr, I. Rabi, “Computer Supported Collaborative Work (CSCW) and Network Issues: A Survey”, *Information Journal Tokyo Japan*, pp. 7995-8020, 2014.
- [30] Zyda, S.S.a.M., “Networked Virtual Environments Design and Implementation”, 1999.
- [31] Spirent Communication, “Spirent Communication White Paper”, Jitter, version 1.00, 2007.
- [32] Gutwin, C. “The Effects of Network Delays on Group Work in Real-Time Groupware in Proceedings of the Seventh European Conference on Computer-Supported Cooperative Work”. Germany: Kluwer Academic Publishers, 2001.
- [33] Vaghi, I., C. Greenhalgh, and Benford. “Coping with Inconsistency due to Network Delays in Collaborative Virtual Environments”. In *Proceedings of the ACM Workshop on Virtual Reality and Software Technology*
- [34] Park, K.a.K., “Effects of Network Characteristics on Human Performance in the Collaborative Virtual Environment”. In *Proceedings of IEEE Virtual Reality ’99*. 1999.
- [35] Bassiouni, M. A. & Mukherjee, A. “Data Compression in real-time distributed Systems”. *Proceedings of Global Telecommunications Conference (GLOBE COM)*, pp. 967-971. San Diego, CA: IEEE, 1990.
- [36] Singhal, S. K, “Effective Remote Modelling in Large-scale Distributed Simulation and visualization environments”, Academic, Department, 1996.

- [37] Singhal, S. K. & Zyda, M. "Networked Virtual Environments New York, ACM Press, 1999.
- [38] <http://ftp.utcluj.ro/pub/users/cemil/prc/CONGESTION%20CONTROL.ppt>.
- [39] <http://www.cs.uni-paderborn.de/fileadmin/Informatik/AG-Karl/Lehre/WS0910/RNetze/cn1-ch7-congestioncontrol-4up.pdf>.
- [40] "Design best practices for latency optimization financial services technical decision maker", 'white paper'.
- [41] Alexey Tumanov, Robert Allison, Wolfgang Stuerzlinger, "Variability-Aware Latency Amelioration in Distributed Environments", IEEE Virtual Reality Conference Charlotte, North Carolina, USA, 2007.
- [42] <http://www.searchnetworking.techtarget.com/definition/propagation-delay>, 11 July 2013
- [43] BERTSEKAS, D. AND GALLAGER, R. "Data Networks". Englewood Cliffs, Prentice Hall, 1992.
- [44] SONG, Y. "Time Constrained Communication over Switched Ethernet". Proceedings IFAC International Conference on Fieldbus Systems and their Application. Nancy, France, 152-169, 2001.
- [45] Anurag, K., Manjunath, D. and Kuri, "Communication Networking: An Analytical Approach". Morgan Kaufmann, San Francisco, 2004.
- [46] <http://www.intel.com/technology/ioaccelation/306517.pdf>
- [47] <http://sagarsharam89.hubpages.com/hub/Optical-fiber-cable-vs-wireless-communication>, 11 July 2013.
- [48] Macedonia, M. R. and Zyda, M. J. "A taxonomy for networked virtual environments". IEEE Multimedia, 4(1):48–56, 1997.
- [49] R. Gossweiler, R. J. Laferriere, M. L. Keller, and R. Pausch. "An introductory tutorial for developing multiuser virtual environments". Presence, 3(4):255{264, 1994.
- [50] S. K. Singhal and D. R. Cheriton. Exploiting position history for efficient remote rendering in networked virtual reality. Presence, 4(2):169{193, 1995.
- [51] Cédric Fleury, Thierry Duval, Valérie Gouranton, Bruno Arnaldi, "Architectures and Mechanisms to efficiently Maintain Consistency in Collaborative Virtual Environments", www.collaviz.org.
- [52] D. Delaney, T. Ward, and S. McLoone. "On Consistency and Network Latency in Distributed Interactive Applications: A Survey – Part II". Presence: Teleoperators and Virtual Env. 15(4):465–482, 2006.
- [53] C. Anthes, P. Heinzlreiter, and J. Volkert. "An adaptive network architecture for close-coupled collaboration in distributed virtual environments". In Proc. of the ACM SIGGRAPH int. conf. on VR continuum and its applications in industry (VRCAI'04), pages 382–385, 2004.
- [54] S. Ullah, M. Millan, "Multi-modal Assistance for Collaborative 3D Interaction: Study and analysis of performance in collaborative work", Thesis report, 26-01-2011.
- [55] Fleury, Cédric, et al. "Architectures and Mechanisms to efficiently maintain Consistency in Collaborative Virtual Environments." Proc. of Software Engineering and Architectures for Realtime Interactive Systems-SEARIS, pp. 87-94, 2010.
- [56] D. Delaney, T. Ward, and S. McLoone. "On Consistency and Network Latency in Distributed Interactive Applications: A Survey – Part II". Presence: Teleoperators and Virtual Env. 15(4):465–482, 2006.
- [57] R. Waters, D. Anderson, J. Barrus, D. Brogan, S. Mckeown, T. Nitta, I. Sterns, and W. Yerazunis. "Diamond Park and Spline: A Social Virtual Reality System with 3D Animation, Spoken Interaction, and Runtime Modifiability". Presence: Teleoperators and Virtual Environments, 6(4):461–480, 1997.
- [58] J. Calvin, A. Dickens, B. Gaines, P. Metzger, D. Miller, and D. Owen. "The SIMNET virtual world architecture". In Proc. of the IEEE Virtual Reality Annual Int. Symp. (VRAIS'93), pages 450–455, Sep 1993.
- [59] C. Shaw and M. Green. "The MR Toolkit Peers Package and Experiment". In Proc. of the IEEE Virtual Reality Annual Int. Symp. (VRAIS'93), pages 463–469, 1993.
- [60] C. El Zammar. "Interactions cooperatives 3D distantes en environnements virtuels gestion des problèmes r'eseau". PhD thesis, INSA de Rennes, 2005.
- [61] M. Mauve, V. Hilt, C. Kuhmunch, and W. Effelsberg. "A general framework and communication protocol for the transmission of interactive media with real-time characteristics". In Proc. of the IEEE Int. Conf. on Multimedia Computing and Systems (ICMCS'99), volume 2, pages 641–646, Jul 1999.
- [62] D. P. Brutzman, M. Zyda, K. Watsen, and M. R. Macedonia. "Virtual Reality Transfer Protocol (VRTP) Design Rationale". In Proc. of the 6th Workshop on Enabling Technologies on Infrastructure for Collaborative Enterprises (WET-ICE '97), pages 179–186, 1997.
- [63] W. Broll. "DWTP - an Internet protocol for shared virtual environments". In Proc. of the 3rd symp on Virtual reality modeling language (VRML'98), pages 49–56, 1998.
- [64] S. Jourdain, J. Forest, C. Mouton, B. Nouailhas, G. Moniot, F. Kolb, S. Chabridon, M. Simatic, Z. Abid, and L. Mallet. "ShareX3D, a scientific collaborative 3D viewer over HTTP". In Proc. of the 13th int. symp. On 3D web technology (Web3D'08), pages 35–41, 2008.
- [65] D. Delaney, T. Ward, and S. McLoone. "On consistency and network latency in distributed interactive applications: A survey – part I". Presence: Teleoperators and Virtual Environments, 15(2):218–234, 2006.
- [66] D. R. Jefferson. "Virtual time". ACM Trans. on Programming Language and Systems, 7(3):404–425, 1985.
- [67] T. A. Funkhouser. "RING: a client-server system for multi-user virtual environments". In Proc. of the symp. On Interactive 3D graphics (SI3D'95), pages 85–93, 1995.
- [68] D. Margery, B. Arnaldi, A. Chauffaut, S. Donikian, and T. Duval. "OpenMASK: Multi-Threaded or Modular

- Animation and Simulation Kernel or Kit: a General Introduction". In Proc. of the Virtual Reality Int. Conf. (VRIC'02), pages 101–110, 2002.
- [69] D. J. Roberts and P. M. Sharkey. "Maximizing concurrency and scalability in a consistent, causal, distributed virtual reality system, whilst minimizing the effect of network delays". In Proc. of the IEEE workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, pages 161–166, Jun 1997.
- [70] S. Jourdain, J. Forest, C. Mouton, B. Nouailhas, G. Moniot, F. Kolb, S. Chabridon, M. Simatic, Z. Abid, and L. Mallet. "ShareX3D, a scientific collaborative 3D viewer over HTTP". In Proc. of the 13th int. symp. On 3D web technology (Web3D'08), pages 35–41, 2008.
- [71] G. Singh, L. Serra, W. Png, A. Wong, and H. Ng. "BrickNet: sharing object behaviors on the Net". In Proc. of the IEEE Virtual Reality Annual Int. Symp. (VRAIS'95), pages 19–25, Mar 1995.
- [72] Zeigler, B. P., Cho, H., Lee, J. S., Cho, Y. K. & Sarjoughian, H. S. "Predictive Contract Methodology and Federation Performance". Proceedings of Fall Simulation Interoperability Workshop (SIW '99) Orlando, Florida, 1999.
- [73] Lin, K.-C. & Schab, D. E. "The performance assessment of the dead reckoning algorithms in DIS". *Simulation* 63(5): 318-325, 1994.
- [74] Capin, T. K., Emerald, J. & Thalmann, D. "A dead reckoning technique for streaming virtual human animation". *IEEE Transactions on Circuits and Systems for Video Technology* 9(3): 411-414, 1999.
- [75] Roberts, D. J., Strassner, J., Worthington, B. G. & Sharkey, P. M. "Influence of the supporting protocol on the latencies induced by concurrency control within large scale multiuser distributed virtual reality system". Proceedings of International Conference on Virtual Worlds and Simulation (VWSIM), SCS Western Multi-conference '99, pp. 70-75, San Francisco, California, 1999.
- [76] Krumm-Heller, A. & Taylor, S. "Using determinism to improve the accuracy of dead reckoning algorithms". Proceedings of SimTecT2000 - The Simulation Technology and Training Conference Sydney, Australia: Simulation Industry Association of Australia, 2000.
- [77] Duncan, T. P. & Gracanin, D. "Pre-Reckoning algorithm for Distributed Virtual Environments". Proceedings of Winter Simulation Conference, pp. 1086-1093, 2003.
- [78] Benford, S. D. & Fahlén, L. E. "Aura, Focus and Awareness". Proceedings of 5th MultiG Workshop Royal Institute of Technology, Stockholm, Sweden: Kungl Tekniska Hogskolan, 1992.
- [79] Benford, S. & Fahlén, L. E. "A Spatial Model of Interaction for Large Virtual Environments". Proceedings of Third European Conference on Computer Supported Cooperative Work (ECSCW '93) (pp. 107-123). Milano, Italy: Kluwer Academic Publishers, Dordrecht, 1993.
- [80] Greenhalgh, C. & Benford, S. "Boundaries Awareness and Interaction in collaborative Virtual Environments". Proceedings of 6th Workshop on Enabling Technologies on Infrastructure for Collaborative Enterprises (WET-ICE '97), pp. 193-198, Cambridge, MA: IEEE, 1997.
- [81] Funkhouser, T. A. RING: "A client-server system for multiuser virtual environments". Proceedings of Symposium on Interactive 3D Graphics (pp. 85-92). Monterey, California, United States: ACM SIGGRAPH, 1995.
- [82] Macedonia, M. R., Zyda, M. J., Pratt, D. R., Barham, P. T. & Zeswitz, S. NPSNET: "A Network Software Architecture for Large Scale Virtual Environments. Presence: Teleoperators and Virtual Environments", 3(4): 265-287, 1994.
- [83] Hook, D. J. V. D., J. C., Newton, M. & Fusco, D. "An Approach to DIS Scalability. Proceedings of 11th Workshop in Standards for the Interoperability of Distributed Simulation", pp. 347-355, 1994.