**Synchronization**

The continuous media streams have well defined temporal relations between the subsequent Media Units (MUs), for example, audio and video streams. The non-continuous media streams like images, text and graphics do not have temporal relations among MUs.

Multimedia presentation requires the integration of multiple media streams of both continuous and noncontinuous streams. These streams have different temporal dependencies among the MUs of one or multiple streams. To ensure these relationships between the MUs of single and/or multiple media streams, a coordination process is required, which is called the multimedia synchronization. Typical synchronization solutions can be classified in to two basic types Intra-media synchronization deals with the reconstruction of the temporal relations between the MUs of the same media stream, at the presentation time. For example, maintaining the frame sequence and frame rate of the video stream to ensure a smooth presentation. During presentation, reconstruction of the temporal relations between the MUs of the different but related media streams is referred as Inter-media synchronization. A typical example of the inter-media synchronization is lip synchronization between the corresponding audio and video stream.

In distributed multimedia environment, apart from the two basic synchronization problems described earlier, another type of synchronization is required in case of multicast communication and is called inter-destination or group synchronization. This is required when geographically scattered group of receivers have to present the same stream(s) approximately at the same. With the emergence of Interactive Distributed Multimedia Applications (IDMA) a new type of interactive synchronization emerges and examples are . In these types of applications, users can modify the presentation state of stream and this modification has to be communicated to all receivers to maintain the synchronized view of the presentation among them.

The current packet switching networks do not provide any guarantee on delay bounds of packet delivery. Rather they only promise best effort to deliver the data to the intended recipient. This characteristic of packet networks make the success of the distributed applications challenging. It causes asynchrony (de-synchronization) in Distributed Multimedia Applications. In the following section, we will briefly discuss the factor of asynchrony.

**Causes of Asynchrony**

MUs of the media stream suffer different type of delays from the generation at source to presentation at receiver. These delays can be different for different MUs depending upon the load at sender, network and receiver. These delay variations for different MUs cause asynchrony in the media presentation at the receiver. We can divide delays into three types: the delay caused by sender, network and by the receiver. Figure1 gives a pictorial representation of all three components of end-to-end delay.

**Delay at sender:** Capturing, coding, packetizing, protocol layer processing and transmission-buffering delays depend on the sender load and clock speed. At the different time instances, the sender may have different loads variations, which can cause the variation in these delays for different MUs. Moreover, if the related sub-streams are captured or/and sent by different sources, then, these delays experienced by different sub-stream can be more variable. **Network delay:**

Network delay is the delay experienced by the MUs in the network to reach its receiver, which varies according to network load. This delay can include the propagation delay and queuing delay at the intermediate routers. Network jitters is delay variations of inter-arrival of MUs at the receiver due to varying network load. This is due to the fact that the queues of the intermediate routers between sender and receiver may have different loads at the different time instances. This delay can cause intra-media asynchrony. Network skew is the time difference in arrival of temporally related MUs of different but related streams, i.e. differential delay among the streams, which can cause inter-media asynchrony. Clock drift is the rate of change of clock skew because of temperature differences or imperfections in crystal clocks. Clock skew is the clock time difference between the sender and the receiver. This is possible if the sender and the receiver are using local clock information instead of global clock information. The sender and receiver are considering time synchronized with respect to clock only if they are using the Network Time Protocol (NTP) or Global Positioning System devices. Figure 1. End-to-end causes of delay.

**Delay at receiver:**

The presentation, decoding, depacketizing, protocol layer processing, and buffering delay at the receivers can be different for different MUs. These delay variations are present at the receiver due to the fact that different receivers may have different processing capabilities and different loads at the different time instance. Depending on the nature of the application some or all of these problems may be relevant to different applications. Different synchronization mechanisms are needed to cope with these problems to ensure the temporal ordering of streams and to maintain the presentation quality.

**INTRA- MEDIA SYNCHRONIZATION**

The reconstruction of temporal relations between media units of the same continuous media stream is referred to as intra-media synchronization. For audio streams, the basic media unit is audio sample. The spacing between samples is determined by the sampling process. The goal of inter-media synchronization is to ensure the same spacing at the presentation time. For video streams, the basic media unit is the video frame and the temporal relation is the frequency of the video frames. The frame rate determines the spacing between the Copyright (c) IARIA, 2012. ISBN: 978-1-61208-195-3 2 MMEDIA 2012 : The Fourth International Conferences on Advances in Multimedia frames. At presentation time, similar frame-rate has to be ensured by reconstructing the temporal relationship. Many schemes have been proposed in literature to ensure the temporal relationship at presentation time. All the schemes use a receiver buffer for the temporary storage of incoming MUs. The audio/video samples/frames are then presented at appropriate time from buffer. The use of a MU buffer introduces delay in the application, which is directly proportional to the size of this buffer. The objective of the process is to provide a presentation that resembles as good as possible to the temporal relations that were created during the encoding process. All Distributed Multimedia Applications (DMAs) have their own end-to-end delay tolerance requirement [33] that depends upon the nature of the application. Interactive bidirectional applications such as online quizzes have very strict end-to-end delay requirements and the applications like video conferencing have slightly less strict latency requirements. On the other hand applications like video on demand (VOD) can allow larger latency. All the proposed schemes provide for a compromise between the intra-media synchronization quality and the increase of end-to-end delay due to the buffering of MUs. On one extreme, there can be a buffer less scheme with minimum delay by presenting the frame as soon as they arrive and other can be assured synchronization that completely eliminate the effect of jitter on the cost of high end-to-end delay. The perfect intra-media synchronization quality can be achieved by completely eliminating any kind of distortion in the temporal relationships of MUs and to completely restore the stream to its initial form. If the delay variability is unbounded, meaning that an infinitely long inter arrival period may appear, then no technique with a finite buffer can eliminate the distortion from the MUs. But, some assured services (QoS) guarantee the bounded network delay. In this case, one can achieve assured/perfect synchronization. We divided the intra-media synchronization in to two basic categories: Time-oriented techniques and bufferoriented techniques. In time-oriented techniques sender puts a time stamp on the MUs. The sender and the receiver use clock in order to measure the delay and jitter. Receiver on the basis of these measurements devises a technique to ensure synchronous presentation of streams. Buffer-oriented techniques do not use the clock rather they implicitly measure network delay and jitter by the occupancy of the receiver buffer. The summary is presented in Table 1. A. Time-oriented Techniques We divide time-oriented techniques into three subcategories, depending upon the timing information: techniques using global clock information, techniques using local clock information, and techniques using approximated clock information. Techniques in which sender and receiver use some mechanism for the synchronization of their clock are said to use global clock information. The existence of having the globally synchronized clock allows the receiver to measure the exact network delay of MUs. Due to exact measurements of network delay, it can guarantee that MUs will be delivered and presented before or at the required time. The techniques “using the global clock information” [7, 8] measure network delays of the first MU. They then add buffering delay in already measured network delay to compose it to total delay. They set the Maximum Delay equal to this total delay. The receiver keeps the first MU in the buffer for minimum of buffering delay time plus the extra interval before extracting from the head of the buffer for presentation. This extra buffering delay for the first MU protects the synchronization of the stream for the succeeding MUs. This way, it is guaranteed that no MU will experience a larger delay then the first MU, thus no loss of synchronization will occur. The amount of this extra buffering delay will decide the quality of synchronization. The larger extra buffering delay means assured synchronization and smaller means small endto-end delay but no assurance of synchronization. The amount of this extra buffering delay can be adjusted according to the nature of the application. For more interactive application this amount can be set low.

**INTER-MEDIA SYNCHRONIZATION**

The inter-media synchronization is concerned with maintaining the temporal and/or logical dependencies among several streams in order to present the data in the same view as they were generated. At the receiver, MUs will not arrive in synchronized manner due to jitter in the network. The temporal relationship within sub-streams is destroyed and the time gaps between arriving MUs vary according to the occurred jitter. Thus, a synchronized presentation cannot be achieved at the receiver, if arriving MUs of sub-streams would be presented immediately. Hence, intra-media and inter-media synchronization is disturbed. To mitigate the effect of the jitter, MUs have to be delayed at the receiver so that, a continuous synchronized presentation can be guaranteed. Consequently, MUs have to be stored in buffer and the size of the buffer will correspond to the amount of jitter in the network. For example, in video conferencing applications speech and video MUs must have the temporal relationships at the time the streams were captured at source. These speech and video MUs captured at the same time have to be presented together at receiver. Like any two different streams, the audio and video stream can be affected by the network delay differently. If these streams would be presented without any synchronization mechanism at the receiver, the audio and the corresponding lip movement in the video will not be synched. This temporal relation between the audio and the video stream is called inter-media synchronization or Lip synchronization. A pictorial representation of lip synchronization is presented in Figure 2. The perfect inter-media synchronization quality is achieved by completely eliminating any kind of distortion in the temporal relationships of MUs among multiple streams and to completely restoring the stream to its initial form. This objective must be achieved on the fly as MUs arrive at the receiver, having crossed a network that alters the spacing between MUs, by imposing a variable network transfer delay.

**INTER-DESTINATION MULTIMEDIA SYNCHRONIZATION (IDMS)**

In multicast media communication, apart from intramedia and inter-media synchronization, we can find another type of synchronization called group or inter-destination media synchronization (IDMS). The objective is to present the same stream at all the receivers in a group, approximately at the same time. To add to the complexity of the problem, these different receivers may be located at different Copyright geographical locations and may have different processing capabilities. These receivers may not only be of different type like smart phone and laptop computer but also may have the network connection of the different speeds. Network quizzes can be a good example of this scenario, where the objective will be to achieve the fairness among all the participants of the quiz. Solution will be required to display all the questions of the quiz to the entire participant at the same time. The other example can be of the real time distance learning (tele-teaching), where the teacher multicasts a multimedia lesson to a number of students, who are located at different geographical areas. In this scenario, the teacher can also make comments about the live streaming of the lesson. Another similar example is of the interactive internet TV (Internet Social TV), where different groups of friends are watching a live online football match at different geographic locations. Consider the case, when these groups can chat (audio/video) to each other to comment on the game to experience of watching the match together from distinct location. It will be very important to synchronize the streams, so that they can watch the different events of the match at the same time to have the real experience of watching together.

Multimedia streams have well defined temporal relations within themselves, generated when captured at the sender. At receiver these temporal relations have to be reconstructed to ensure smooth and synchronized multimedia presentation.

The characteristics of best effort network –delay and jitter- degrade the temporal relations present in multimedia streams.

Distributed multimedia applications like video conferencing, video on demand, distance learning and others, are made feasible due to developments in the communication network. In such applications, at sender's side, different media streams are captured and sent to the receiver via packet switching network. On the receiver side, streams are received for presentation.

These media streams can be classified into continuous and non-continuous streams. The continuous media streams have well defined temporal relations between the subsequent Media Units (MUs), for example, audio and video streams. The non-continuous media streams like images, text and graphics do not have temporal relations among MUs.