Ce document a été élaboré dans le cadre des travaux de veille technologique (Vtech) de la Plate-forme THD

Fiche "Audio-Video"

- **Auteur :** Sandoche BALAKRICHENAN (AFNIC)

- **Dernière mise à jour :** 17/12/2009
- **Diffusion :** "Portail / Tout-Public"

As per Cisco nearly 90% of consumer IP traffic is expected to consist of video on demand, IP peer-to-peer video, and Internet video in 2012. All these types of traffic fall under the same category: real-time multimedia transmission. The objective of the THD project is to set up a platform where new digital services based on high speed networks are experimented and tested. Majority of these digital services will involve real-time multimedia applications. Real-time multimedia applications are basically audio and video being transmitted and played back in real-time.

Before we talk about the current developments in audio and video transmissions, there is a need to recall how the network architecture has been evolving. At the beginning (around 40 years ago) only voice was transmitted using the well-known telephone network (also called as Public Switched Telephone Network (PSTN)). When data transfer became a necessity IP (Internet Protocol) networks were used. So initially voice was transmitted through the telephone networks and data were transmitted through the IP networks each separately. Then came the demand for providing video over IP services. Managing separate networks for voice (over PSTN) and data/video (over IP) led to high operational and maintenance cost. The focus then moved on to have a single IP network wherein voice, video and data could be transmitted. This is normally mentioned as "triple-play" in the operators terminology.

For an end-to-end service (can be voice, data or video packets) that is from the source to the destination, the information could pass through two different types of networks: Core and Access networks. The core network is used for the transmission of data at the national or global level. The data is then passed to the access network for distribution at the destination end which is usually termed as the "last mile".

There is a variety of access technologies such as xDSL (Digital Subscriber Line) which is more commonly used at present to newer technologies such as FTTH (Fiber-To-The-Home) or FTTC (Fiber-To-The-Curb) delivering up to 100 Mbps to the subscriber.

TCP (Transmission Control Protocol, RFC 793) and UDP (User Datagram Protocol, RFC 768) are the common transmission protocols used in IP networks. The advantage of UDP is that it does not add much overhead to the transmitted packet. This is possible because UDP does not take into account flow control and reliability. This is also its disadvantage. With UDP, it is difficult to check if the transmitted packet has reached it destination or not (typically lost in the way). So there arises the arguement whether to use TCP over IP. TCP makes sure that the packet transmitted reaches its destination and in proper sequential order. But these advantages of TCP have a cost: overhead and also latency.

For multimedia applications, timelines of the packet reaching its destination is a very important criterion. For example, latency of more than 250 milliseconds between two sequential voice packets can be detected by human beings. Hence TCP is not a preferred option for a large number of multimedia applications.

Then it leaves only UDP, but UDP does not guarantee that real-time (e.g. video conference, or hearing a song online) data will reach the destination properly (e.g. the packets received at the destination are not in the same order as sent or they are lost or did not arrive within the stipulated time). Hence there is a necessity of a new protocol that makes sure that multimedia data is received with correct synchronization and timing.

RTP (Real-Time Transport Protocol, RFC 3550) is one such protocol that is built[1] on UDP and that provides timestamping, sequence numbering, and other mechanisms to take care of the timing issues. Through these mechanisms, RTP provides end-to-end transmission for real-time data over UDP. RTCP (Real-Time Transport Control Protocol, RFC 3550) is the control protocol that is designed to work in conjunction with RTP. During an RTP session, RTCP provides feedback to the sender on the quality of data distribution. The sender can adjust its transmission based on this feedback. RTP and RTCP were developed at the IETF Audio-Video transport working group (`http://www.ietf.org/dyn/wg/charter/avt-charter.html`).

It is clear that the the current Internet architecture by default does not provide the same physical wire security as the phone lines. For example during a VoIP (Voice over IP session), it is quite possible that an intruder can tamper with the voice data. The key to securing multimedia packets during a real-time transmission is to use the security mechanisms like those deployed in data networks (such as firewalls, encryption, etc.). The SRTP (The Secure Real-time Transport Protocol, RFC3711) provides confidentiality, message authentication, and replay protection to both RTP and RTCP.

SRTP or SRTCP (Secured Real-time Transport Control Protocol) works on the principle of public and private key infrastructure. This protocol assumes that both the sender and the receiver have a known public key and a private key. SRTP provides authentica-

[1] From design point, RTP is designed to be independent of the underlying protocols

tion of the RTP payload and header. Algorithms used in SRTP could not only provide authentication but also integrity of the source.

Most of the multimedia content has been transferred following RTP over UDP stream format. Innovations such as in the rapid synchronization (as provided by RTP/RTCP) could be a necessity even while sending through faster channels such as FTTH. Adoption of content protection for Video on Demand or games online is a must, hence secured transfer of the RTP payload/header (as provided by SRTP/SRTCP) becomes obligatory.