TITLE

[0001] Multiple Connections for Real-Time media

DESCRIPTION

[0002] As required, detailed embodiments of the present invention are disclosed herein; however, it is to be understood that the disclosed embodiments are merely exemplary of the invention that may be embodied in various and alternative forms. The figures are not necessarily to scale; some features may be exaggerated or minimized to show details of particular components. Therefore, specific structural and functional details disclosed herein are not to be interpreted as limiting, but merely as a representative basis for teaching one skilled in the art to variously employ the present invention.

[0003] One non-limiting aspect of the present invention relates to a method for increasing the available bandwidth and improving the reliability for real-time media sessions between two media endpoints by utilizing multiple connections between the endpoints.

[0004] IP networks may become the defacto standard transport for real-time media exchanged between user endpoint devices. Real-time media over IP networks has its challenges, however. The main problem is that the network is a shared resource that can become congestion, causing packets to be delayed or lost. Most of the congestion occurs in the access network – the segment of the network that provides connectivity between the user's endpoint device and the internet service provider's backbone IP network.

[0005] Service Provider's have implemented QoS mechanisms to help resolve these congestion issues (e.g., PacketCable Multimedia™ provides dynamic QoS for real-time media over DOCSIS®), but there are still many situations where a user wants to establish real-time communications and the IP network does not have sufficient resources to provide good quality.

[0006] A media endpoint can have multiple IP addresses, where each IP address is obtained from a different physical access network. For example, a mobile phone could have two IP addresses;

one for cellular radio access via 4G network and one for Wi-Fi. Or, a SIP-PBX could have three IP addresses; each obtained over a separate DOCSIS Cable Modem.

[0007] While contemplated, one non-limiting aspect the present invention is not referring to the case where an endpoint obtains multiple IP addresses on a single access network; e.g., a dual-stack endpoint that obtains an IPv4 address and multiple IPv6 addresses all on the same access network.

[0008] When an endpoint with multiple access networks wants to establish a real-time media session with a remote endpoint, the present invention contemplates it setting up a single connection over one of its access networks to carry the media (in other words, the cardinality of media session to IP connection is 1-to-1). One non-loony aspect of the present invention contemplates that instead of using only a single access network for the session, the endpoint uses multiple access networks simultaneously. Basically, the endpoint establishes multiple connections via multiple access networks to the remote endpoint, and sends media across all connections.

[0009] This can provide two benefits:

[0010] **More bandwidth:** If there isn't sufficient bandwidth in any single access network to support the media type, then the endpoint can establish multiple connections with the remote endpoint – one connection per access network – and distributing the media across the multiple connections to provide the required bandwidth (the available bandwidth being the sum of the bandwidth of each connection).

[0011] **Improved reliability:** There are cases where any one of the available access networks has sufficient bandwidth to carry the target media type most of the time, but each access networks experiences random intervals of high packet loss. In this case the user can improve the quality of the call by establishing multiple connections – one per access network – and sending the same set of media packets on each connection. Assuming that the periods of packet loss are randomly distributed across the connections, the probability of delivering at least one of the media packets for each sample interval is greatly increased.

[0012] The mechanism contemplated by the present invention may assume that the endpoints have knowledge of the bandwidth and packet-loss characteristics of each local access network. At call setup time, the local and remote endpoints negotiate the number of connections to be established, and how those connections are to be used in order to increase bandwidth and/or reliability. The following list describes the different use-cases:

[0013] **Case-1**

[0014] Near Endpoint: connects via multi access networks for more bandwidth (say, traffic split 75/25 between IP1 and IP2)

[0015] Far Endpoint: connects via a single access network (say from IP3). Far Endpoint creates an access pipe that fits 100% of traffic. It accepts downstream traffic from IP1 & 2, and distributes upstream traffic to IP1 & 2 at negotiated ratio 75/25.

[0016] IP1 — 75% --> IP3

[0017] IP2 — 25% --> IP3

[0018] IP1 <-- 75% -- IP3

[0019] IP2 <-- 25% -- IP3

[0020] **Case-2**

[0021] Near Endpoint: connects via multi access networks for better reliability (all packets sent via IP1 and IP2).

[0022] Far Endpoint: connects via a single access network (say from IP3). Far endpoint creates an access pipe sized to carry 2X the number of packets of a single connection (i.e., 200%). It accepts downstream traffic from IP1 & IP2, discarding redundant packets. It sends each upstream packet twice; once to IP1 and once to IP2.

[0023] Near endpoint discards redundant received packets.

[0024] IP1 — 100% --> IP3

[0025] IP2 — 100% --> IP3

[0026] IP1 <-- 100% -- IP3

[0027] IP2 <-- 100% -- IP3

[0028] **Case-3**

[0029] Near Endpoint: connects via multi access networks for more bandwidth

[0030] Far Endpoint: connects via a multi access networks for more bandwidth

[0031] Since both endpoints need multiple connections for more bandwidth, they negotiate the number of connections and the ratio of media sent on each connection. Say in this case they negotiate 2 connections, from near-end IP1&2 to far end IP3&4.

[0032] IP1 — 75% --> IP3

[0033] IP2 — 25% --> IP4

[0034] IP1 <-- 75% -- IP3

[0035] IP2 <-- 25% -- IP4

[0036] **Case-4**

[0037] Near Endpoint: connects via multi access networks for more reliability

[0038] Far Endpoint: connects via a multi access networks for more reliability

[0039] Since both endpoints need multiple connections for more reliability, they negotiate the number of connections, and send all packets on all connections. Say in this case they negotiate 2 connections, from near-end IP1&2 to far end IP3&4.

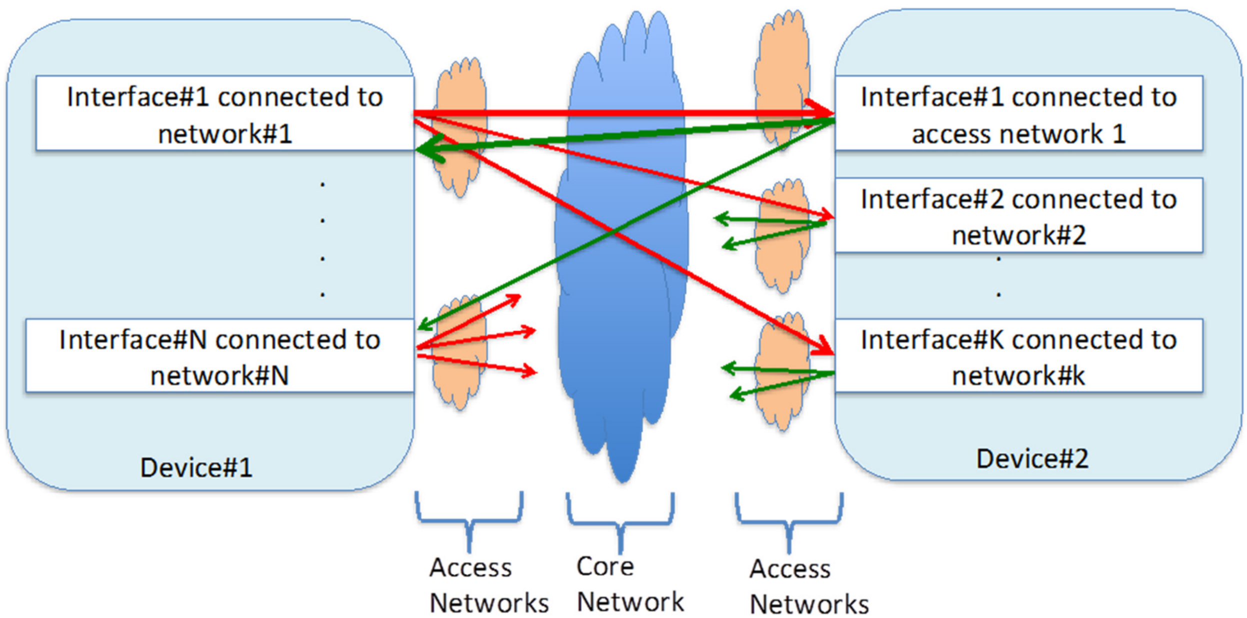
[0040] IP1 — 100% --> IP3

[0041] IP2 — 100% --> IP4

[0042] IP1 <-- 100% -- IP3

[0043] IP2 <--100% -- IP4

[0044] The following diagram shows a more general example for Case-4, where both the local and remote endpoints negotiate multiple connections and the portion of media sent/received on each connection.



- Arrows between devices show multiple paths used for a real time media session
- Thickness of the line indicates different amount of media packets on different paths
- Paths are selected based on various factors such as: bandwidth, cost, managed vs un-managed, real time feedback for each path (e.g. using extended RTCP),

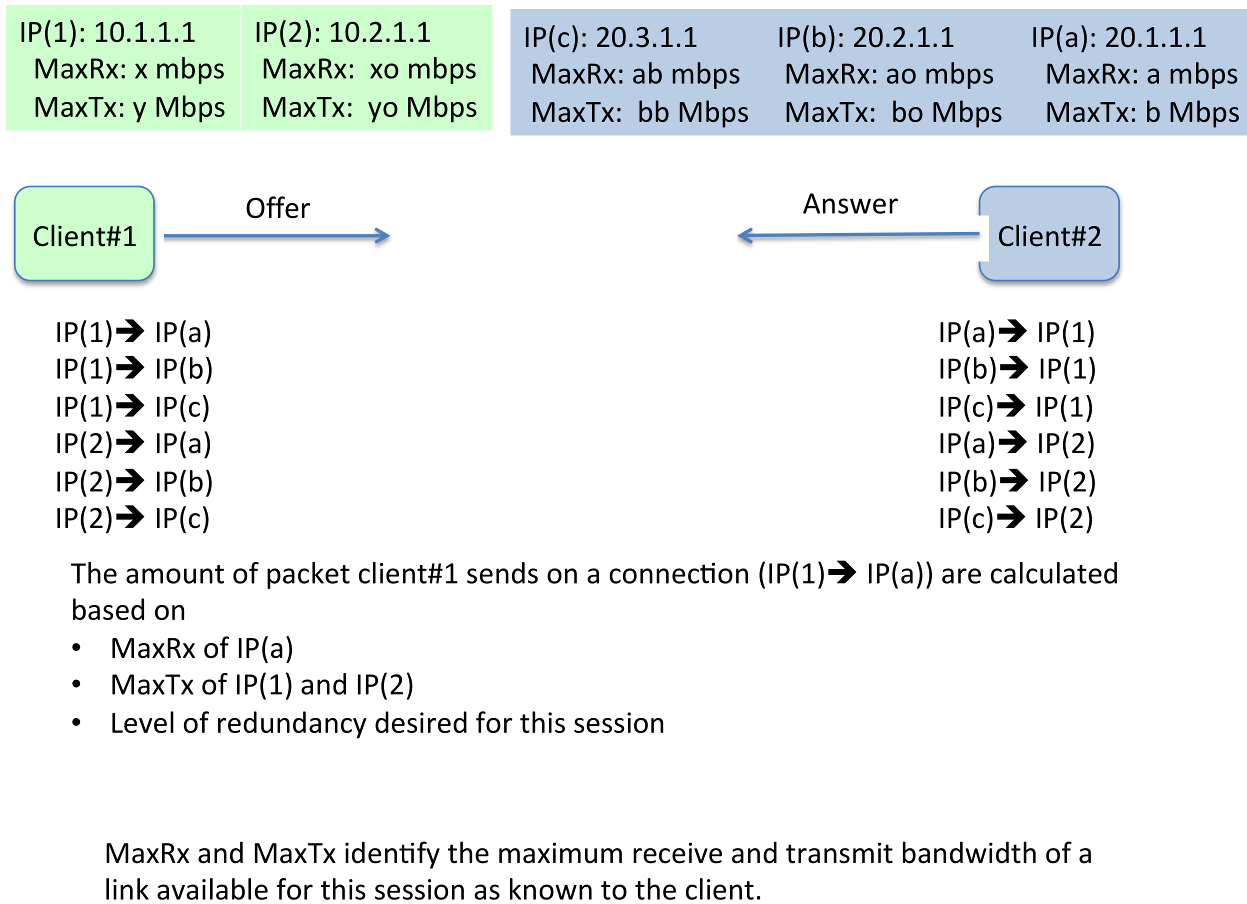
[0045] **Case-5**

[0046] Near Endpoint: connects via multi access networks for more bandwidth

[0047] Far Endpoint: connects via a multi access networks for more reliability

[0048] In this case the endpoints could negotiate a combination of M-to-N local-to-remote access network connections, and also negotiate the media distributed (including redundant media) sent/received on those connections.

[0049] The following diagram shows an example of Case-5.



[0050] This idea can further be extended to use real time link utilization information, where a client is capable of leaning status (e.g. how busy) of each link during the active session and make changes to the connection used for the session. For example, the changes can include moving some of the data from link that is busier to the ones that is less busy.

[0051] Additional details:

[0052] Our disclosure talks about selection of both outgoing interface (on transmitting device) and incoming interface (on receiving device) by the transmitting device. For example, in our disclosure a transmitting device takes the following things into consideration before transmitting packets:

[0053] * Rate at which packets will be generated by the application on the transmitting device

[0054] * Bandwidth supported by various outgoing interface on the transmitting devices

[0055] Bandwidth supported by various incoming interfaces on the receiving device

[0056] * Size of packets (since some links are better suited for smaller packets and others are more suited for large packets)

[0057] * Delay characteristics of each link (if the RTCP packets are sent over high delay link, there may not be any adverse affect to customer experience)

[0058] If there are 3 IP interfaces (TX1, TX2, TX3) on the transmitting device and 3 IP interfaces (RX1, RX2, RX3) on the receiving device, then after taking into consideration the list provided above, the transmitting device chooses the pair of outgoing link (on TX device) and incoming link (on RX device) to use for this transmission. For example, the transmitting device may end up selecting:

[0059] TX1 ==> RX2 (for 20% of the traffic generated by the application)

[0060] TX2 ==> RX3 (for 30% of the traffic generated by the application)

[0061] TX3==> RX 1 (for 20% of the traffic generated by the application)

[0062] TX3==> RX 2 (for 30% of the traffic generated by the application)

[0063] Algorithms to choose Access Networks:

[0064] A client (client#1) interested in establishing a real time media session with another client (client#2). Client#1 sends a session establishment request (e.g. SIP INVITE) with an offer. The offer from client#1 will include the following information:

[0065] * All the IP addresses that Client#1 has, including IP address for each interface

[0066] * Information about all the Access Network that client#1 is directly connected to

[0067] * Known characteristics of each Access network. Some example of the characteristics include:

[0068] - Downstream Bandwidth

[0069] - Delay, jitter, packet loss

[0070] - MTU

[0071] - Price

[0072] * The client#2 receives the offer and replies back with an answer. The answer can include the following information:

[0073] * All the IP addresses that Client#2 has, including IP address for each interface

[0074] * Information about all the Access Network that client#2 is directly connected to

[0075] * Known characteristics of each Access network. Some example of the characteristics include:

[0076] - Downstream Bandwidth

[0077] - Delay, jitter, packet loss

[0078] - MTU

[0079] - Price

[0080] At this point each client not only knows about information available locally but also the information from the other end. Each client takes into consideration the information available locally and from the remote end to decide how many packets should be sent on each source and destination IP address combination. Client is not only paying attention to available bandwidth but also delay and jitter characteristics to decide which application (e.g. audio vs video) packets within the session should be send on what link.

[0081] The invention takes advantage of the multiple access network technologies available today (especially wireless networks) to improve the reliability and throughput of real-time media transmitted over IP networks.

[0082] For example, say an endpoint with two access networks – 3G and Wi-Fi – wants to establish a hi-definition video call. The endpoint knows that there isn't sufficient bandwidth available on either the 3G or the Wi-Fi network to support the hi-def codec, but that the total available bandwidth across both networks is sufficient to support the codec. Therefore, it establishes two connections to the remote endpoint, and distributes the media traffic across both connections. Without this mechanism, the endpoint would have been unable to establish the call using the target codec.

[0083] Or, say a user has three access networks; a Wi-Fi network whose throughput is affected by sporadic electromagnet interference, a 3G network which randomly experiences bursts of high packet-loss due to temporary overload conditions, and a 4G network whose throughput is occasionally diminished due to temporary increases in multi-path distortion. Rather than the endpoint having to select one of the access networks to establish a real-time media connection, with the result that the call has periods of poor quality. By using the new mechanism described here, the endpoint can establish multiple connections and send/receive the same media on each connection, this greatly decreasing the probability if lost packets and greatly increasing call quality.

[0084] While exemplary embodiments are described above, it is not intended that these embodiments describe all possible forms of the invention. Rather, the words used in the specification are words of description rather than limitation, and it is understood that various changes may be made without departing from the spirit and scope of the invention. Additionally, the features of various implementing embodiments may be combined to form further embodiments of the invention.

[0085] Today's wireless access networks are plagued by poor quality due to oversubscription and environmental interference of various kinds. As a result, mobile endpoints are forced to use extreme compression algorithms in order to fit real-time media into the available access network bandwidth. The end result is that users must put up with poor quality. This invention disclosure enables the endpoint to obtain more bandwidth by adding on the bandwidth resources of multiple access networks. As a result the endpoint can use higher-quality codecs, thus providing a better user experience. Tthe ability to decrease packet loss through redundancy across multiple networks can enable the use of codecs that have poor packet-loss concealment algorithms but are otherwise good (e.g., royalty-free, very high quality/definition).