Packet Aggregation - overcoming packet-per-second limits

PROBLEM

Every communications device will have its limits. Some modems with integrated routers will have a limit of performance that becomes particularly apparent when running VoIP traffic. This is due to the increased number of small packets generated by VoIP phones directly through the modem.

SOLUTION

Vocality can provide a suite of functions to overcome the packet-per-second (PPS) limitation experienced over satellite modems on non-encrypted links. It achieves this through a packet aggregation function within the Vocality device.

In addition, to ensure the network still benefits from voice optimization, the Vocality unit will address the requirement to optimize VoIP traffic, through the use of its PACE (Packet Aggregation for Communications Efficiency) set of functions, including:

- Header compression for RTP traffic
- SIP SDP 64k codec filtering
- Jitter buffering for RTP traffic
- G729B small sample filtering

As shown in the diagram above, a network optimized in this way requires a Vocality device to be installed behind the modem in the remote site, with many of them connecting via single Vocality hub units sited behind the hub satellite modem bank.

PACE Technical Details

Packet Aggregation

Vocality uses packet aggregation, putting many smaller packets into one packet, across its WAN aggregate link. In this way, the satellite modem has fewer small packets and instead handles a number of larger packets. This minimizes the IP overhead required. The immediate benefit is the ability to support a significant number of additional voice channels, when combined with header compression, which would now need to be provided by the Vocality device rather than the satellite modem.

Header Compression for RTP traffic

Portions of the IP, UDP & RTP headers are constant throughout an RTP session (conversations). The Vocality unit can find these conversations, inform the IP peer of the contents of the headers, and then send only a conversation identifier. When the compressed packet arrives at the target IP peer, the headers are reconstituted and sent to the RTP target. RTP sequence and timestamp information is forwarded intact. So each RTP packet appears on the aggregate with 32 fewer bytes. For G729k packets, this is a saving of 52%.

Each packet to be sent to the peer is examined to identify valid packets (IP & UDP checksums) and known conversations. Any UDP traffic with an even port number (except the SIP port number 5060) is considered RTP traffic. Service management filters ensure that only RTP port numbers used in the target network are forwarded down the RTP compression enabled channel.

SIP SDP 64k codec filtering

The SIP SDP 64k codec filtering logic looks for session description protocol messages within SIP signalling messages. It strips out any PCMU and PCMA (G711 64k) codec negotiation to prevent SIP devices selecting 64k codecs to use over the network. Note: This relies on the SIP signalling stream running non-secured over UDP. It also only works for ASCII format SDP messages.

Jitter Buffering for RTP Traffic

Some SIP devices cannot tolerate the large jitter seen in some IP satellite networks. The Vocality unit seeks to remove this jitter from an RTP stream that is forwarded through the embedded IP router and across the satellite network. RTP packets that are forwarded from the IP router to an IP trib are pre-pended with a timestamp. This timestamp is sent across the satellite network with each RTP packet, and the timestamp is used at the peer unit to forward packets onward at the same frequency with which packets arrived. Note: This relies on timestamps from the CPU clocks on peer units across the Vocality network and using these to control packet synchronization.

G729B Small Sample filtering

When a G729B codec sends data over an RTP stream it typically generates 20ms samples of 20 bytes (plus RTP/UDP/IP overhead). During "normal" conversations some G729B devices generate smaller samples (of 10ms) as voice/silence transitions occur. Several of these may occur per second during "normal" speech. If these voice samples are not forwarded, the perceived voice quality is not greatly affected but the bandwidth required to forward them is saved.