Planning Networks for VOIP

An Introduction

Contents

1 Introduction

Setting up a network to carry Voice over IP (VOIP) traffic along with the data packets requires an understanding of what features are needed in the routers and switches of the network. It is also helpful to understand how much data bandwidth is required to successfully transport VOIP voice conversations, and what tradeoffs are possible to get by with less bandwidth than would normally be necessary.

In this paper, we will try to give some advice in planning a network that will carry voice and data traffic with a minimum of problems. We will look at voice quality requirements, network layout, planning capacity, and router selection and configuration.

2 Voice Quality Requirements

Users have certain quality expectations for telephone communication. When they are talking to a customer through a VOIP link, for instance, they might insist on toll-quality voice. This means that they want it to sound to them and to the customer as though they are on an ordinary TDM telephone link.

The quality of a call can be measured using one of several call quality metric calculations. The most commonly used system is the Mean Opinion Score (MOS). The MOS score of a call is between 1 (for unusable) and 5 (for excellent). VOIP calls that are working properly fall between 3.5 and 4.2. 4.0 is defined as toll quality. Other systems for quality measurement are R-factor, PSQM, PESQ, and PAMS. These other systems produce scores for a call that can be mapped to MOS for comparison.

MOS score is an indication of what users would think about the call. It was developed using surveys of users of different technologies, but today it is calculated through the use of engineering formulae.

MOS	Listening Quality	*Listening Effort*
5	Excellent	Complete Relaxation
4	Good	Attention Necessary
3	Fair	Moderate Effort
2	Poor	Considerable Effort
1	Bad	No Meaning Understood

When designing an IP network to carry voice, we cannot allow the MOS to drop below 3.0 at any time. For most networks, it should be above 3.7 at all times.

3 Codecs

The VOIP device (VOIP telephone, gateway, PBX, key system, or adapter card) will
encode voice into a packet format for the IP network, and another device will decode the
data back into voice at the other end. The name of this Coding-Decoding process is
shortened to "Codec". There are different codecs available to use for connecting voice
calls from one VOIP endpoint to another. These different encoding-decoding methods
have different call qualities and different bandwidth requirements. Basically, if you want
your calls to use less network bandwidth, then you will suffer reduced call quality.

Codec	ms/packet	Bandwidth (Kbps)	*MOS*
G.711	20	83	4.1
G.726	20	58	4.0
G.729	20	26.4	3.92
G.723.1	30	18.4	3.65

The low-bit-rate codecs use a lot less bandwidth and will save you money on your
network links. However, your users will be able to tell that the G.729 calls are "tinny"
compared to the G.711 calls. Also, there will be a problem if the same call gets encoded
and decoded several times between one end and the other because the negative effects on
the audio quality accumulate with every compression cycle.

For example, if a user calls into the office on a G.729 encoded link from a VOIP phone at
home, connects through a G.729 encoded tie-line to a branch office, and then goes out an
outside line to connect to a person with a GSM encoded cell phone, the overall call
quality will be an unacceptable 2.68 MOS. This is due to the degradation in the quality
of the call each time it is decoded and encoded.

The G.711 PCM codec does not have this problem. Since it has one-to-one mapping with
the TDM PCM in the phone network, there is no loss of quality due to encoding and
decoding. The G726 ADPCM algorithm also does not suffer from these types of quality
loss due to repeated encodings and decodings.

Also, G.729 tends to exaggerate any network problems like jitter and packet loss. The
3.92 MOS shown is for ideal network conditions, but this degrades exponential when
faced with a WAN that has problems. It will move into the unacceptable sub-3.0 range
very quickly if packets reach their destination late or fail to arrive at all.

Later, when we discuss capacity planning, you will see where you have to decide
between saving some money by choosing G.729 or assuring excellent call quality by
choosing G.711.

4 Network Layout

A typical company data network has high-speed LAN switches and routers in each location with lower-speed connections between locations. Often, each office has a single connection to an ISP and that ISP has a high-speed backbone network that connects each of its locations to one another.

Figure 1 shows a typical layout for a company with three locations and VOIP equipment in each location.
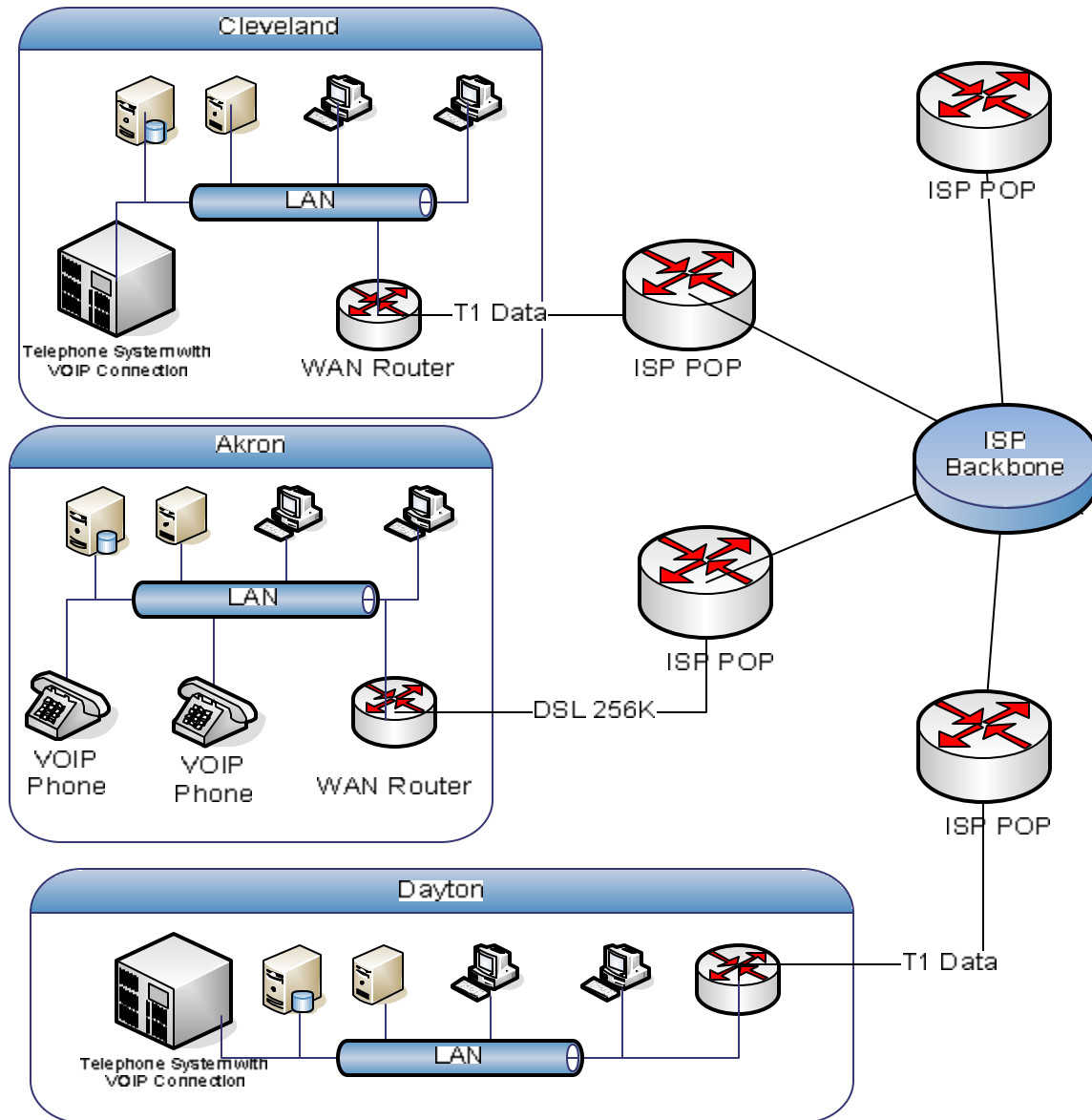

Figure 1 - Typical Customer Layout

In this example, each location includes a WAN router that connects to a router at the ISP through a lower-speed link   In Dayton and Cleveland this link is a T1, while in Akron it is a DSL line. These links could also use other technologies like microwave, point-to-point wireless, ISDN BRI, or even dialup.

The most important thing to note about this network is that the customer has installed a QOS capable router at each location. This router must be capable of providing LLQ-type queuing for the VOIP packets, or there will be serious problems with the voice quality of the calls that go through the link. Additionally, there must be a network engineer to configure the router who understands voice QOS requirements. This cannot be overstated. If the customer does not have the expertise to do this configuration, then he should lease the router from the ISP and verify that the ISP can configure and maintain the QOS features of the router for him.

5   Planning Capacity

There is a fairly straightforward method for planning the amount of bandwidth that you will need to carry your voice and data traffic over your lower-speed WAN links. We will go through it step-by-step.

5.1 Step 1 – Estimating Voice Bandwidth

The voice traffic and its associated call control packets will be marked and handled separately from the data traffic (See "What VOIP needs from the Data Network"). The amount of bandwidth required for voice is the number of calls times the amount of bandwidth per call. The bandwidth per call is a factor of the codec and the packetization interval (in ms.)

Codec	ms/packet	Bandwidth (Kbps per call)
G.711	20	83
G.711	30	76
G.726 (32K)	20	58
G.726 (32K)	30	51
G.729	20	26.4
G.729	40	17.2

So, in our example, there are two VOIP phones in Akron using the G.729 codec with a packetization interval of 20 ms. If both phones are in use at once, the voice requirement of the WAN connection to the ISP would be 26.4 X 2 or 52.8 Kbps. In Dayton, the telephone system can support eight G.729 calls over the WAN at 20 ms./packet, so the Dayton WAN connection should be engineered to support 26.4 X 8 or 211.2 Kbps of voice traffic.

We should add 10% to this estimate to cover the call control traffic that will also be specially marked for Low Latency Queuing. In our example, we are planning to use G.729 with 20 ms./packet for all of our connections.

Location	Calls	Bandwidth Per Call	Voice-only Bandwidth	Total Voice (incl. 10% call control)
Akron	2	26.4 Kbps	52.8 Kbps	58 Kbps
Dayton	8	26.4 Kbps	211.2 Kbps	233 Kbps
Cleveland	8	26.4 Kbps	211.2 Kbps	233 Kbps

5.2 Step 2 – Estimating Data Bandwidth

It is really beyond the scope of this document to describe how to estimate data bandwidth requirements for your WAN links. This requires an understanding of how many users will be at each location, where the applications servers are in relation to the users of those servers, how much bandwidth each user of each application needs, etc.

If the routers are configured properly, then data will be limited so that it does not take bandwidth away from voice. However, when all of the voice lines are in use, there is some bandwidth that is required to maintain your data applications.

5.3 Step 3 – Under subscription

When planning data networks, network engineers put a self-imposed limit of 75% of the bandwidth specified by their provider. This gives a little extra room to accommodate the fact that data usage tends to come in bursts. Users choose to download files at random times. It is the nature of random events that they seem to occur in bunches. Several users will download files at once, followed by a much lower usage for a while. If the data usage were uniform throughout the day, then the events would not be random.

If the network is primarily used for data, then we add 30% to the estimated total bandwidth usage to account for this phenomenon.

5.4 Step 4 – Calculating Total Required Bandwidth

When we configure the routers, we will classify voice traffic and call control traffic as Low Latency for LLQ treatment. This means that the voice packets will always take priority over other packets. You should not assign more than one-third of your WAN bandwidth to LLQ.

For most installations, you can calculate the total voice bandwidth and multiply by three to find your total required link bandwidth. This assumes that your required data bandwidth is less than twice the necessary voice bandwidth.

We have added some Total Data requirements to our example:

Location	Total Voice	Total Data	3 X Voice	Voice+Data+30%
Akron	58 Kbps	128 Kbps	174 Kbps	242 Kbps
Dayton	233 Kbps	256 Kbps	700 Kbps	636 Kbps
Cleveland	233 Kbps	768 Kbps	700 Kbps	1301 Kbps

Note that the voice+data requirement for Dayton is only 636 Kbps, but we must still get a WAN connection that is more than 700 Kbps. This is because our LLQ voice of 233 Kbps should not exceed one third of the total connection. After doing this analysis, we know that we need at least the larger of 3 X Voice and Voice+Data+30% for the WAN connection to each location.

5.5 Step 5 – Compromises

If you cannot afford to connect the WAN link that this analysis requires, you must sacrifice some voice quality to fit within the WAN bandwidth that you have.

- If you chose G.711 for your codec in Step 1, try running the same calculations with G.729. This will have a small, predictable impact on call quality

- If you have equipment that supports 8 calls, see if you can limit it to 4 calls or 6 calls.

- If lowering your data traffic would lower your bandwidth requirement (as it would in Cleveland in our example), see if you can move data servers to make this possible.

You should never try to force more VOIP voice traffic through the link without LLQ protection. This may work in some limited tests, but it is likely to fail unpredictably on the busiest day of the year.

For example, let's see what would happen if we connected Dayton to the ISP with a 512Kbps partial T1. According to our calculations, Dayton needs to be able to make 8 simultaneous VOIP calls and do 256Kbps of data.

We could configure the router to use one third of the 512K bandwidth for LLQ priority voice. That would be 170Kbps, which would support about 5.5 G.729 calls. This would seem to work fine until there were 6 or more calls active at the same time. Then the router would queue some of the voice packets behind data packets, and they would arrive too late. This would cause dropouts in all 6 calls. The actual quality would not be predictable. The users might be able to live with it, or they might find the quality so poor that it makes the connection unusable.

Alternatively, we might think that we could boost the size of the LLQ in the router to one half the bandwidth of the connection or 256Kbps. This would probably carry all of the voice calls without quality problems, but the data traffic would suffer. When there were

eight calls going at once, the router would queue all of the voice packets ahead of the data packets. There would not be enough bandwidth left for the data applications to run smoothly.

6   Router Selection and Configuration

6.1 Router Requirements

The most important steps toward good voice quality in your VOIP network are choosing and configuring the routers that connect your VOIP equipment to the WAN. The routers must support priority queuing for the voice packets, so that they can go out ahead of the data packets.

The routers that need this configuration most are the WAN routers at the edge of each campus and the POP router at the ISP's location. However, any router or switch that lies on a path between the two VOIP devices should support priority queuing. Look for switches that have "Quality of Service" features in their description, and routers that support "Low Latency Queuing". Replace any Ethernet hubs in the path that voice traffic uses with switches.

6.2 Router Configuration

To configure the routing, you will need an engineer who is well-qualified to deal with VOIP Quality of Service issues. He will need to configure the routers and switches to put all of the VOIP voice and call control packets into a Low Latency class that is handled appropriately. The Low Latency class must be set to go into the Low Latency Queue (LLQ) or Priority Queue (PQ) so that it goes out of the switch or router ahead of all data traffic.

Additionally, features of the router like Weighted Random Early Detection (WRED) that drop packets before they go into the queue should never be used for voice packets. Dropping packets early can cause unexpected variations in latency (jitter) that will cause voice communication to fail.

Finally, some method of fragmenting large data packets like LFI is very helpful for slow WAN links. Any link slower than 768Kbps will have trouble carrying large (1500 byte) data packets and small voice packets. The time it takes to serialize and send a single 1500 byte email or file download packet will have a negative impact on the voice packet that is next in line. To help solve this problem, LFI breaks the large data packets up into fragments and will interleave voice packets in between the fragments.

6.3 Traffic Shaping Appliances

Instead of configuring a sophisticated router to perform these functions, some enterprises buy a traffic shaping device that goes between the LAN switches and the WAN router. These devices, like ones from Packeteer, will queue and reorder packets according to the LLQ strategy and do the fragmentation and interleaving of large data packets. This takes some of the processing burden off the WAN router, and can lead to improved latency.

7 The Big Picture

The objective in data network planning is to create a data network that will carry the traffic with delays that users can live with. The objective in creating a voice network is to build one that will support voice conversations that users will consider toll-quality. Designing a converged voice-data network requires you to meet both of these objectives.

You can do that by:

❖ Specifying routers, switches, traffic shapers, and ISP partners that will support the QOS features described in this paper.

❖ Having or finding the expertise necessary to configure that equipment

❖ Planning bandwidth capacity that is the larger of:

➢ Three times the Voice over IP bandwidth

➢ Voice + Data + 30%

❖ Choosing Codecs that meet your voice quality needs while fitting within the bandwidth budget