Due to the many questions I have been asked about VoIP and Digital Lines with Fax and Fax Servers, I have assembled the following information from our vendors, industry experts and my own experiences.

The state of technology is always moving so some enterprising genius could have this problem solved tomorrow. As of today, the following statements may help you when approached on this issue.

Questions? Email is your best bet. admin@lesolson.com

- Keith

Some Basic "true-isms" with FoIP (Faxing over IP)

Sending FAXes over VoIP networks usually fails. It is human nature to look for simple reasons for that, and simple cures. In reality, there are a number of reasons, and no certain universal cures. VoIP networks are designed to do a good job with speech. Carrying any sound other than a single voice speaking is not generally a system requirement. It shouldn't be too surprising if it works rather poorly.

Modems don't like relativity:

In the PSTN (Public Switched Telephone Network) world, the network provides a constant delay for any particular call. The speed at which data enters the network is always the same as the speed at which it leaves. The end to end delay does not jitter, or make step changes in anything but exceptional circumstances (e.g. on automatic fail-over, if a fibre link fails). Modems require this. In an IP network jitter is a fact of life. It can be kept to a modest level, through the use of the QoS (quality of service) features available in a lot of IP equipment, but only if you control the network end-to-end. If the call passes across the open Internet there is no QoS control. It is hard to see a business model that would ever encourage QoS to be introduced across the open Internet. So, in the long term the timing of a voice signal entering a VoIP network is the same as the timing as it leaves, but in the short term they can be very different.

If a VoIP network works only across a LAN or a QoS managed WAN link, there might be a near guarantee of zero packet loss, and fairly low jitter. Many people then assume the jitter buffer at the receiver will smooth out the modest jitter, and the received signal will perfectly match the transmitted one. They are often right, but there is no guarantee. There are many designs of jitter buffer. Most modern ones dynamically adapt the length of the buffer in some way, although many different algorithms are used. If the jitter is low, and dynamic jitter buffering is switched off, things may work well. If it cannot be switched off, the behavior of dynamic buffering will generally upset a modem signal. Various algorithms will:

- Guarantee some packet loss, by tuning the buffering until a small percentage of packets are declared late, and dropped. Dropping packets is actually built into these algorithms, and the results can be pretty good for voice. Trading a small number of dropped packets for somewhat less latency is a reasonable trade-off.
- Adjust periods of silence in whole packet steps (typically 20ms). Certain silence periods in a FAX signal are specified as 75+-20ms. 20ms jumps can push them out of spec.
- Continuously adjust the timing of non-silent periods, using overlap and add techniques. This is the state of the art in jitter buffering for voice, but a complete disaster for a modem.

The more basic equipment is likely to work well, and the newer more sophisticated designs are likely to be troublesome.

Wide Load, narrow road:

The most common problem with sending a FAX over VoIP networks is the easiest to deal with. A low bit rate voice codec is unable to carry a fast modem signal without severe distortion. Would you really expect an 8kbps G.729 codec to convey a 9.6kbps FAX modem signal correctly? The only common codecs capable of adequately preserving FAX modem signals up to 14,400bps (V.17) are u-law and A-law. Up to 9600bps (V.29) a fully implemented G.726 codec will also work. However not all codecs claiming to be G.726 fully implement the spec. A few shortcuts can save considerable compute power, and only a few people need the spec. to be fully implemented. The G.726 codec was, however, specifically designed to be able to carry medium speed modems, such as the V.29 modem used for FAX.

FAX machines supporting 33,600bps (V.34bis) have become popular. This rate is unlikely to work with any reliability across any VoIP connection, even when an A-law or u-law codec is used. The codecs will maintain the required signal quality, but the delay across the VoIP channel, even if it is a stable delay, will prevent the echo cancelers in most modems from training well enough. The slower FAX modems - V.27ter, V.29 and V.17 - do not use echo cancellation, so the problem does not exist there.

Lower bit rate codecs have zero chance of working for any standard FAX image modem. Many will convey the 300bps (V.21) FAX control messages OK. They will not convey the fast modem signals, used for the actual image data.

Modems don't like silence suppression:

Depending on its implementation in particular equipment, silence suppression can destroy a FAX call. If silence suppression is enabled, a voice detector continuously monitors the call, looking for the presence of a real voice. Some of these are designed to focus purely on voice, and tend to reject other kinds of sound - e.g. modem tones. They may, therefore, not switch the audio path on and off cleanly when the modem signal starts and stops. Even if they do switch cleanly, the suppression algorithms usually modify the audio around the switching points.

During silent periods, comfort noise is usually introduced, to simulate the background noise you normally hear in a conversation. This might mean a period which should be silent, is actually significantly noisy. The receiving modem might not see a good enough "silence" for its signal detector to correctly declare the boundaries of the modem signal.

Conclusion:
FAX, and other modem applications, operating over VoIP channels are quirky, and unreliable. This will **not** get better over time. It will get **worse**. In general, the more sophisticated the equipment gets in trying to make speech work smoothly, the worse it behaves for modems. In the near term (i.e. until all data applications are native IP applications) store and forward protocols, and protocols tailored to reasonably conveying modem data across an IP channel are the only way to achieve consistent results.

This portion of the document is a compilation of data from a VoIP Codec design expert supplemented with my comments about dedicated G3/G4 modem Fax technologies. The design expert is with a company called AudioCodes which is credited for the G.723 VoIP compression standards. His remarks are in **Blue**. My inserted comments are in **Orange**

The main parameter that will never become standard is the preservation of voice quality when moving from TDM to IP. TDM networks are based on G.711 coding, which makes all voice equal or with the most negligible differences. This is where the fax is designed to work within the TDM or Time Division Multiplexing environment over PSTN (Public Switched Telephone Networks). An IP network is never the same because of the number of factors influencing the ability of a network to move a packet from one destination to another.

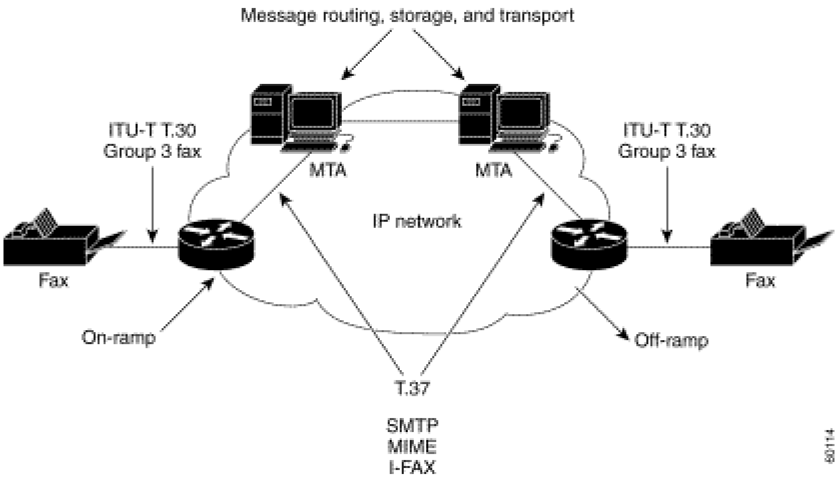
Even the same IP network will not provide the same quality of service at any given point of time. The environment is constantly changing. The ability to maintain premium quality while moving voice on top of changing networks depends on the way the media gateway's mechanisms are designed to work to overcome all of these factors. As well as the hardware responsible for transport over a QoS designed network for VoIP. Such as routers, Switches, and whether the IP network crosses over VPN or changes framing types. None of this work is standardized. Years of work must be invested in designing algorithms to manage voice transfer as best as possible.

While a call is in progress, the challenge is to compress without losing voice quality. Or in this case, the ability for a consistent signal from one G3 device to reach another G3 device with the same timing and delivery rates as provided by TDM over a PSTN. The mechanism for compression differs from one media gateway to another, with the compression algorithms being the variable.

One challenge for media gateway makers is to create buffers to handle jitter, or delay variation. If delay were at a constant pace, the other party would not notice. But in VoIP delay is a changing factor. Buffering streamlines the packets. If there is no buffering the sound gets to the other side with different time variations and won't sound right. The buffer streamlines the packet transfer and arrival. This can be affected by packet loss or MTU's as well There can be different delay variations during a conversation.

Designing an optimal buffer is a challenge for a media gateway provider. Small buffers are cheaper to make, but won't provide the tool to ensure quality. Causing packet loss across a transmission which will result in a fax error A buffer that is too large can be the source of delay. Causing timing issues for the receiving fax and also resulting in a fax error. Therefore the challenge is to adjust the size of the buffer to solve delay without creating delay. Larger buffers will make a product more expensive but can manage changing delay situations. This issue can be addressed using a "Store and Forward fax Gateway"

One method to bypass traditional Fax issues over VoIP is to use a "Store and Forward Gateway". This system takes the traditional G3 type fax signal receives the fax converts it to an email message and then forwards the email to another store and forward gateway where it is converted back to fax and delivers it to a G3 fax machine.



Another expertise is packet loss concealment, or the technology for covering up for lost packets. Fax transmissions can not accept concealed packet loss. Each packet is an element of the finished product. Voice packet loss concealment is acceptable because the human ear/mind combination will fill in gaps automatically, especially when they only last a fraction of a second.

Algorithms can be created in the media gateway to fill up the lost packets with something to cover for the cutoffs, which are usually only a fraction of a second. The goal is to predict the lost packet and fill it in.

Double talk is another challenge to maintain when two parties are talking together to ensure that there is no speech clipping and discontinuity when two parties talk at the same time. The echo canceller suppresses residual echo to the background noise level so it will be perceived by the user as regular background noise rather than disturbing echo.

Achieving satisfying VoIP quality is not easy, and will always require ongoing, and significant R&D investment, especially in proprietary algorithms. In fact, maintaining cutting-edge voice quality in the media gateway is all about R&D.

All media gateways are therefore not born equal and as yet, have not eliminated voice quality issues throughout their evolution. Also, since the media gateway is only a single element of the topology, you must also take into account Routers, Switches and line quality along the IP network.

Often times a system (especially in a large environment) may initially function with Fax and errors may be less frequent, however since the IP network is ever changing and in nearly never dedicated to voice only (for cost reasons) will eventually change. Often it is the addition of a feature or service to the VoIP system that precipitates the change.

Much investment must be made in creating features to preserve state-of-the-art voice quality, the most important factor in a voice network. Only when this happens, will voice quality become an issue of the past.

This is how a Traditional Fax (PSTN) Call Flows: