Voice over Internet Protocol

Design Considerations, Guidelines, and Practices

Table of Contents

Introduction

Intended Audience

This publication is intended for all Inter-Tel employees and authorized providers who sell, install, and service Inter-Tel's various Voice over Internet Protocol (VoIP) products and services, including:

- Salespeople
- Sales Engineers
- Field Technicians

Purpose

The publication provides a brief overview of various topics associated with VoIP and explains why these topics are important in achieving good quality VoIP calls. Although there are now many detailed books and references materials on the topic of VoIP, there are few, if any, that cover the basic VoIP topics in a simple, compact format similar to the one found in this publication.

The information in this publication is not a substitute for formal data networking and VoIP training and certification, such as those found in Network+ and Convergence Technology Professional (CTP) certification programs.

By bringing together many relevant VoIP topics under one cover, this publication can act as a general VoIP reference guide or be used to refresh knowledge on a specific topic or topics. Should the reader require detailed information beyond what is outlined in this publication, there are numerous resources available in bookstores, on the Internet, and in training programs.

Format

Each topic in this publication is divided into four main sections:

- Definition
- Measurement
- Tuning
- Effects

Each topic is kept intentionally short (generally one to three pages) to allow the reader to get some basic knowledge in less than 10 to 20 minutes.

The reader is introduced to the topic with a brief overview or definition. Next is an explanation of which measurements, if any, can be made. The tuning section explains what, if anything, can be configured. And, finally, the effects of tuning are addressed.

When reading this document, Inter-Tel strongly recommends that the reader make use of complimentary information found in other publications, such as:

- Inter-Tel Technical Reference Manuals
- Inter-Tel Knowledge Bases
- Inter-Tel University Training and Courseware
- Industry Standard Specifications such as TSB 116, ITU G.711, ITU G.729, EIA/TIA-464, etc.
- Industry Standard Certification Programs
- Processes defined by Technical Support and the System Sustaining Group
- Technical Books
- The Internet

Corrections

For the latest available revision of this document, see the Online Manuals & Guides section on the edGe. Although every reasonable effort has been made to ensure that the information contained in this publication is accurate and up-to-date, there could be errors or omissions. Please report all errors, issues, concerns, feedback, etc. to Inter-Tel's Technical Publications department at: Tech_Pubs@inter-tel.com.

In addition to corrections, Inter-Tel welcomes any feedback or improvement suggestions. It is our intention to update this document on a periodic basis to add more topics and improve existing ones.

Acknowledgements

This publication was put together as a volunteer effort on the part of about 10 employees from Inter-Tel Integrated Systems in Chandler who are considered Subject Matter Experts (SMEs) on one or more of these topics. These volunteers encompass such disciplines as engineering, support, training, data, audio, and technical writing. Despite full workloads, these people volunteered to work on this project because they knew how important this information was to ensuring the successful deployment of VoIP.

Audio Delay

Definition

Audio delay is the time required for speech to travel from the microphone on one endpoint to the speaker on another endpoint. For VoIP, delay usually refers to a constant or slowly-changing delay. This should not be confused with jitter, which is the quick change in delay. The major factors contributing to audio delay are vocoder delay, packetization delay, network delay, and jitter buffers.

Vocoder delay is the time required to collect, encode, and decode a full frame of audio. Vocoders typically output audio frames with a fixed time interval. For example, G.729 outputs 10 ms frames. The processor or DSP also requires time to encode and decode the frame.

Packetization delay is the time that the first frame in the RTP packet must wait before the endpoint sends the packet. The source endpoint combines several vocoder frames into larger RTP packets, which it sends out on fixed intervals. Most Inter-Tel IP devices use 30 ms RTP packets. For example, with a 10 ms vocoder delay and 30 ms RTP packets, the packetization delay is 20 ms.

Network delay is the delay induced by the data network. Data networks take time to transport packets from one host to another. This network delay is typically determined by link speed and the number of router hops. Router congestion also affects network delay but typically shows up as jitter rather than consistent network delay.

Jitter buffers intentionally add delay on the receiving endpoint to compensate for jitter and provide a consistent stream of packets to the decoder. Static jitter buffers attempt to maintain a fixed delay. For example, the default setting for an Inter-Tel IPRC is an 80 ms jitter buffer. Dynamic or adaptive jitter buffers adjust the delay as required based on the estimated jitter. Many of the Inter-Tel IP phones have dynamic jitter buffers.

Measurement

Vocoder, packetization, and jitter buffer delays depend on the selected vocoder, packet size, and jitter buffer size. Many endpoints have diagnostics displays to show the current jitter buffer size.

Many tools are available to estimate network delay, but the simplest is the common ping utility. Running ping across a network will estimate the average round trip delay. Dividing the average round trip delay by two results in an estimated network delay. This is not an exact measurement, but is sufficient for most VoIP applications. Networks that implement QoS and give low priority to ICMP packets may make measurements using the ping utility less accurate.

To estimate the total peer-to-peer delay, calculate the sum of the vocoder, packetization, network, and jitter buffer delays. A typical peer-to-peer call with G.729, 30 ms RTP packets, 100 ms round-trip ping times, and an 80 ms static jitter buffer has an approximate end-to-end delay

of 160 ms. IP calls which are not peer-to-peer have a vocoder, packetization, network, and jitter buffer delay for each hop.

Tuning

The first steps to reduce audio delay are usually reducing the jitter buffer and packet size as much as possible. If that does not provide desirable results, reducing network delay usually requires an increased link speed or a reduction in the number of router hops. See the VoIP Data Network Requirements document for information about recommended network delays.

Effects

Echo and talk-over are problems typically associated with audio delay. Delay does not cause or amplify echo, but it can make existing echo less tolerable. With a very small delay, less than 5 ms, echo sounds like normal side tone. As delay increases, the echo becomes more noticeable and annoying to the person speaking. Talk-over occurs when speakers frequently interrupt each other because the delay makes it difficult to determine when the other end is finished speaking.

Attempting to reduce delay by adjusting the jitter buffer can reduce the endpoint's ability to manage jitter. If an endpoint's jitter buffer is not adequate, the endpoint will experience packet loss and reduced audio quality.

Most Inter-Tel devices allow packet sizes ranging from 10 ms to 80 ms which is adjustable with the Frames Per Packet setting. Reducing the packet size will decrease packetization delay at the expense of increasing bandwidth requirements.

Bandwidth

Definition

Bandwidth is the amount of data that can be transmitted in a fixed amount of time. For digital devices, bandwidth is usually expressed in bits per second (bps) or bytes per second (Bps).

Bandwidth is often referred to as a data pipeline. When the amount of data traffic attempting to use the pipeline is less than the total capacity of the pipe, data enters the pipe unimpeded. However, when the amount of data traffic attempting to use the pipeline exceeds the capacity of the pipe, there is a bottleneck that can impede entry into the pipeline. In either case, data flows quickly and smoothly once in the pipeline.

Bandwidth is often confused with speed. Increasing bandwidth (enlarging the pipeline) does not increase data transmission speed. Once in the pipeline, all data travels at the same speed.

Measurement

To determine whether you have sufficient bandwidth, you need to know how much of the total available data transfer capacity is being consumed by all of the devices (not just VoIP devices) that send and receive data through the pipeline. If there are multiple pipelines in a host-to-host transmission, the smallest pipeline in the path determines the overall bandwidth capacity.

For good quality VoIP calls, the available network capacity should be approximately 32 kbps per call using the G.729 vocoder and approximately 86 kbps per call using the G.711 vocoder. These numbers are based on a 30 ms packetization interval.

Bandwidth utilization formulas for VoIP devices are generally based on items like selected vocoder, audio frames per IP packet, and RTP profile. For example, see the bandwidth utilization formulas outlined in the latest IP devices manual.

Tuning

If you do not have enough available bandwidth (utilization exceeds capacity), you can increase the bandwidth (enlarge the data pipeline), use Quality of Service (QoS) to give priority to VoIP calls, or tune network data devices so that they consume less of the existing bandwidth. (Note that attempting to adjust device utilization can negatively affect audio quality.)

Effects

Insufficient bandwidth can have a direct effect on jitter, packet loss, and latency (delay), which in turn can result in poor VoIP quality. The only negative of too much bandwidth is paying for unused capacity. However, bandwidth needs often grow over time, so it is a good idea to plan for the future. For VoIP, having excess bandwidth is better than not having enough.

Data Network Security

Definition

The data network includes all devices on the network, including, but not limited to, servers, PCs, routers, switches, and VoIP devices. Note that VoIP devices are members of the data network, but VoIP is not a specific concern.

The primary data network security concern is that no device should allow unauthorized users access to the data network or to compromise systems on the data network. Specifically, VoIP devices should not compromise the data network security. No system, including Inter-Tel devices, should bridge internal and external networks unless specifically designed to do so.

Tools

Various tools can be used to prevent unauthorized access or usage of the data network, including the following:

- Passwords
- Firewalls
- Virtual Private Networks (VPNs)

Change all passwords regularly to unique non-default values that follow generally acceptable password complexity rules (i.e. alphanumeric values that include upper and lower cases and are not dictionary terms).

Firewalls are applications or devices that prevent unauthorized communications forbidden by the network policy. A network firewall limits communication between networks or network segments. Personal firewalls limit communication between the network and a single device.

A Virtual Private Network is a private communications network that communicates over the public network. A VPN can be used to limit access to the data network to select users/devices. A VPN can also encrypt all inter-network traffic to prevent unauthorized parties from obtaining information concerning traffic traveling within the VPN.

Tuning

VoIP audio and signaling streams may not be able to pass through more than one firewall. This must be taken into account when designing the network. Depending on the firewall product (network or personal, hardware or application, etc), the firewall may influence the network delay.

VoIP devices may not support the necessary encryption to support usage on a VPN. This must be taken into account when designing the network. Depending on the VPN scheme, encryption type,

HW or SW encryption, etc., the VPN may influence the network latency (delay) and the network load (bandwidth) as well as impose higher network infrastructure requirements (cost).

Note: If the Inter-Tel system were to sit outside the network, port 80 can be turned off so that the system does not respond to port scanners.

Effects

Adding data network protection tools, including those mentioned above, can add to the network overhead. This can add to the network latency (delay), the network load (bandwidth), and add to the network infrastructure requirements (cost).

Disclaimer

Although no telecommunications system or data network is entirely secure, as long as the appropriate security measures are put in place and properly maintained by both the customer and the installing company (including the proper implementation of user/administrative accounts, passwords, firewalls, NAT, virus protection, security updates, etc. and the proper maintenance of access points/programs and their respective accounts/passwords), the Axxess system architecture and its associated server-based applications are substantially secure against unauthorized access to the customer's data network via the telecommunications system.

DTMF Encoding

Definition

DTMF (Dual Tone Multi-Frequency) is the technical acronym for push button or touchtone dialing. Two tones are mixed together to form a dual-tone. The presence of a specific dual-tone indicates a specific digit or character from the phone's touch pad.

DTMF dialing has been around for many years and is transmitted over standard PCM (G.711) circuit switched connections. DTMF digits become an issue for VoIP calls because VoIP calls can introduced audio impairments (such as distortion due to compression and/or packet loss) that degrade the DTMF digits to the point where they may not be recognized at all, or not recognized correctly (e.g., double digits), by a DTMF receiver.

Because there are multiple encoding choices for sending DTMF over VoIP, such as RFC 2833, G.711, etc., it can be confusing as to which choice is the right choice for specific scenarios.

Applicable standards that address DTMF include, but are not limited to the following:

- EIA/TIA-464A – *PBX Switching Equipment for Voiceband Application*
- RFC 2833 - *RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals*

For a more in-depth discussion of understanding and troubleshooting DTMF, see the document entitled "Understanding DTMF Encoding for VoIP Connections" found on *the edGe*.

Measurements

When encountering problems with DTMF digit transmission, there are two tests points:

- The data network (e.g., measure packet loss)
- The analog side (e.g., measuring the actual analog tones being sent)

DTMF digit corruption can occur due to packet loss on the network. For a discussion of network measurements and packet loss, please refer to the section that discusses that topic.

For measurements in the analog domain, such as over phone lines, special test sets can be used to inject and analyze DTMF tones. These test sets range in price from a few hundred dollars to thousands of dollars depending on the features and functionality you need. Some devices simply monitor and display digits detected. Other devices will provide you with exact measurements of DTMF tone characteristics such as frequency, amplitude, timing, and twist so that you may determine whether the digits are within "spec." A quick search on Google or any other major Internet search engine for "DTMF Test Set" or similar phrasing will provide you with a wealth of choices.

The specs that you will want to measure were listed in EIA/TIA-464A (which has since been updated to "C") under section 4.6.2 ("Dual Tone Multifrequency. Here's a quick summary of the main measurements for DTMF receivers (as opposed to transmitters) that are important to know. DTMF receivers must recognize DTMF signals . . .

- as short as 40 milliseconds (msec)
- -25 to 0 dBm amplitude per frequency
- +4 to -8 dB twist
- Extraneous frequencies must be at least 16 dB below the DTMF frequencies

There are certainly many more measurements besides those listed, but when it comes to problems associated with DTMF digits not being properly transmitted and detected, it's usually because of one of three problems:

1. The digits were distorted (which results in higher harmonics & extraneous frequencies)
2. The timing was distorted (which results in digits shorter than 40 msec)
3. The levels were modified (usually attenuated) – which may result in levels being out of spec.

Tuning

There are two areas you can "tune" with regard to DTMF:

1. Which type of DTMF encoding you ***prefer*** for each IP device.

   Note: When setting the DTMF encoding type for a device, you are simply selecting its preferred encoding type. The actual encoding type used is determined on a call-by-call basis through negotiation with the other device at the start of the call. If you need to determine the actual encoding type used on a call, this can be difficult to do. You may need to use Ethereal and snoop on the packets unless the endpoint tells you this information. For more details about this topic, see the document entitled "Understanding DTMF Encoding for VoIP Connections" found on *the edGe*.

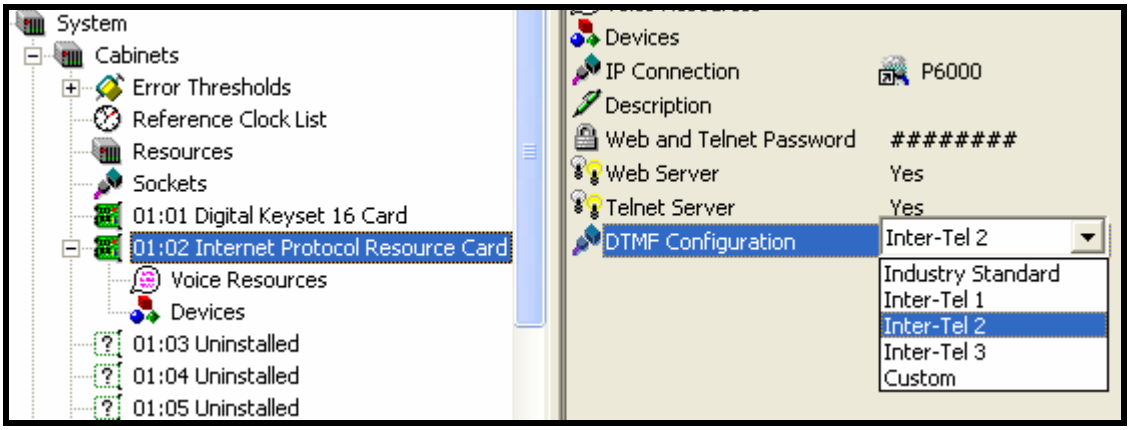
2. In some cases, you may be able to configure or adjust some DTMF parameters for a specific IP device

We essentially have two choices when it comes to deciding on what DTMF encoding scheme we prefer for each VoIP device:

1. Inband audio tones (e.g., G.711, "Same as Voice", "Transparent DTMF", etc.)
2. Named Telephone Events (NTEs) - (using RFC 2833)

There are pros and cons associated with each encoding scheme, but in general, RFC 2833 is the preferred DTMF encoding scheme for VoIP calls because it is more immune to problems caused by packet loss.

Besides choosing the preferred DTMF encoding scheme, some devices will also allow you to tweak or configure the DTMF transmitter and/or receiver to make it more immune to the problem of "talkoff." Talkoff refers to the problem where voice is erroneously detected as a DTMF digit. This can result in spurious DTMF tones being generated in the middle of speech. For example, the Inter-Tel Axxess IPRC will allow you to select from different DTMF configurations (via the OnLine Monitor) to minimize talkoff and maximize digit detection (see the screenshot below).



In general, you won't normally need to adjust the DTMF configuration as the default setting should work sufficiently well.

Effects

The main problems you will encounter with DTMF digits are the following:

- Missed digits (usually because digits are out of "spec")
- Double digits (could be due to encoding technique combined with lost packets)
- Talkoff (impossible to completely eliminate, but "tuning" DTMF parameters can help minimize)

DTMF Receivers can have two types of performance issues: talkoff and missed digits. The ideal DTMF receiver has no talkoff and never misses a valid digit. In reality, every DTMF receiver has some talkoff and sometimes misses digits. It is a tradeoff between one or the other. That is, the more sensitive a DTMF receiver is to DTMF tones, then the more sensitive the DTMF receiver will be to talkoff, because DTMF tones are, by default, a subset of speech.

Talkoff is when a DTMF Receiver detects speech as a DTMF tone. This is undesirable. No DTMF receiver is 100% immune to talkoff, but some DTMF receivers do have higher immunity to talkoff than others.

Suggestions for Trouble-Free DTMF signaling over VoIP

1. Whenever possible, select RFC 2833 as the preferred DTMF encoding scheme for a device unless there are performance issues with the DTMF receiver performance (such as talkoff or missed digits). **Note:** RFC 2833 incorporates the need for DTMF receivers and has the inherent problems associated with DTMF receivers.
2. If you are going to send DTMF digits as audio tones over IP, then always use a high-quality vocoder such as G.711. **Note:** You must have very low packet loss (i.e., high fidelity) to be able to use G.711 as the preferred DTMF encoding scheme.
3. Avoid sending DTMF digits as audio tones over IP using a high-compression vocoder such as G.729 as the higher compression distorts the tones enough that it may cause unreliable detection of the digits.

Ethernet

Definition

Ethernet is a standard protocol for local area network (LAN) communications. As a LAN protocol, Ethernet can carry other *network* protocols including IP, IPX/SPX, DECnet, etc. Ethernet is the de facto standard for VoIP communications on LANs.

The original Ethernet specification, as defined in IEEE 802.3 in 1976, described a 10 Mbps LAN protocol that operated on coaxial media. It was designed to utilize a *bus* topology in which all connected devices shared the same piece of coaxial cable sharing a single receive channel and a single transmit channel. All devices can listen at the same time, but the devices have to take turns transmitting. Because the devices take turns transmitting, Ethernet is considered *half-duplex*. The rules that determine who is allowed to transmit are known *media access control* (MAC). VoIP quality can be impaired when devices need to wait to transmit.

The Ethernet standards have been expanded significantly in recent years to include higher bandwidths (100 Mbps, 1 Gbps, 10 Gbps), different media ("thinwire" coax, unshielded twisted pair, fiber optics, wireless), and functional extensions (VLANs, QoS, full-duplex, Power-over-Ethernet), but the basic operation of Ethernet remains essentially unchanged.

Ethernet information is grouped into *frames* addressed with source and destination MAC addresses (not to be confused with IP addresses). Each Ethernet device is identified by a unique (in the world) MAC address. Each Ethernet device listens to *all* frames on a network until it sees a frame with its MAC address. All other frames, except broadcasts, are ignored. Although all devices listen simultaneously, only one device may *transmit* at a time. Ethernet utilizes a media access control protocol known as *Carrier Sense Multiple Access/Collision Detection* (CSMA/CD).

When a device has a frame to transmit, it waits until no other device is transmitting. When the media becomes available (nobody else is transmitting), the device attempts to transmit its frame. If two or more devices happened to be waiting, they would attempt to transmit at the same time causing a *collision*. When a collision occurs, each of the devices involved stops transmitting and waits a random amount of time before attempting to transmit again.

Although collisions are a normal part of Ethernet's media access control, collisions can increase dramatically in a busy network. Excessive collisions effectively reduce the usable bandwidth in an Ethernet network resulting in delay, jitter, and lost data which collectively degrade VoIP audio quality. The set of devices that compete for media access and potentially cause collisions is known as a *collision domain*. As the number of devices in a collision domain increases, so does the probability of delays, collisions, and lost frames.

Ethernet's media access control protocol, CSMA/CD, was created to coordinate the transmissions of many devices in a collision domain such as when using coaxial cable or a hub. In modern networks, switches divide a LAN into many small collisions domains, usually consisting of just the switch itself and a single device. Besides greatly reducing the likelihood of

collisions, switches can alleviate the need for media access control. When there are only two devices on a network, the first device's transmit channel can be the second device's receive channel and vice-versa. In this scenario, the media access control is not necessary, and both devices can transmit at the same time, also known as *full-duplex*. Since full-duplex Ethernet communications utilize a different protocol (no media access control), full-duplex must be negotiated between the switch and the device or manually configured in both places.

Measurement

The communications on an Ethernet network are best measured using a LAN protocol analyzer (e.g. Ethereal, Sniffer, etc.) Protocol analyzers capture *all* frames on the media regardless of MAC address. Note that using a protocol analyzer in a switched environment is more difficult because switches are specifically designed to segregate the traffic according to destination MAC address. It may be necessary to utilize a hub or a switch feature known as *port mirroring* to allow a protocol analyzer to collect the appropriate information.

The important measurements on an Ethernet network include network utilization and collisions. Many Ethernet devices themselves, including switches, provide counters of frames sent and received, collisions, etc.

Tuning

If the measurements indicate LAN congestion, the first step in tuning is usually to identify where hubs are used and replace them with switches. In nearly all applications, switches will perform significantly better than hubs.

In some cases, significant performance improvements can be gained by utilizing full-duplex connections, particularly to server devices (including VoIP servers). A switched, full duplex Ethernet network usually provides acceptable performance for VoIP, but if necessary, some Ethernet devices support QoS frame prioritization using the IEEE 802.1p standard.

Effects

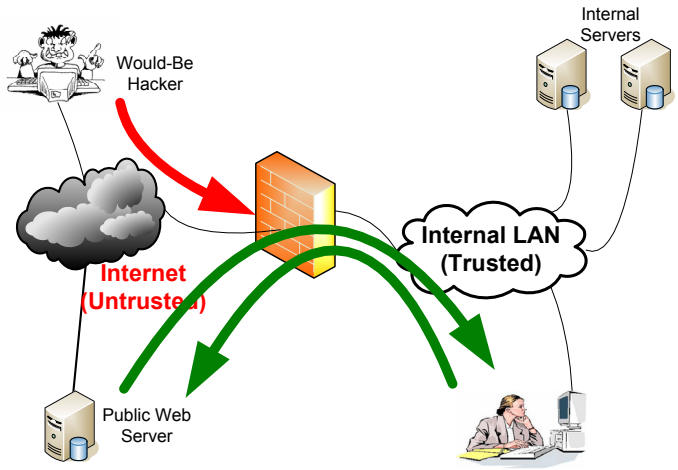
Ethernet (LAN) congestion can cause significant impairments to VoIP. However, Ethernet bandwidth, typically 100Mbps, is relatively large compared with VoIP bandwidth requirements (typically <=64kbps). As a result, well-designed switched Ethernet environments are usually suitable for VoIP traffic.

Firewalls

Definition

A firewall is a set of related programs that provides the connection between a private trusted internal network and the Internet. A firewall is essentially a specialized router, but it may consist of hardware, software, or both. All networks connected to the Internet have, or should have, a firewall in place. Firewalls sometimes can't automatically distinguish between legitimate VoIP calls and security threats. Therefore, firewalls often require specific configuration to allow VoIP.

There are many ways of implementing a firewall, but the main purpose of all firewalls is to protect a trusted internal network from the security threats on the open Internet. A firewall will prevent communications initiated from the Internet from entering the internal network. However, to allow internal users to access the Internet, a firewall must allow communications to go *out* to the Internet and the corresponding responses to return back to the internal network.



Consider the example of an internal user surfing the World Wide Web. The firewall prevents anybody on the Internet from initiating a connection to an internal computer. However, if an internal user types in a URL in their browser, the response from the web site is allowed to come back to the user's browser. In other words, most firewalls operate by keeping track of outbound requests and temporarily allowing the anticipated responses. Unfortunately, VoIP communications, particularly voice streams, do not typically follow this internal-request/external-response structure. As a result, firewalls often block VoIP communications because it may appear to be initiated from the Internet.

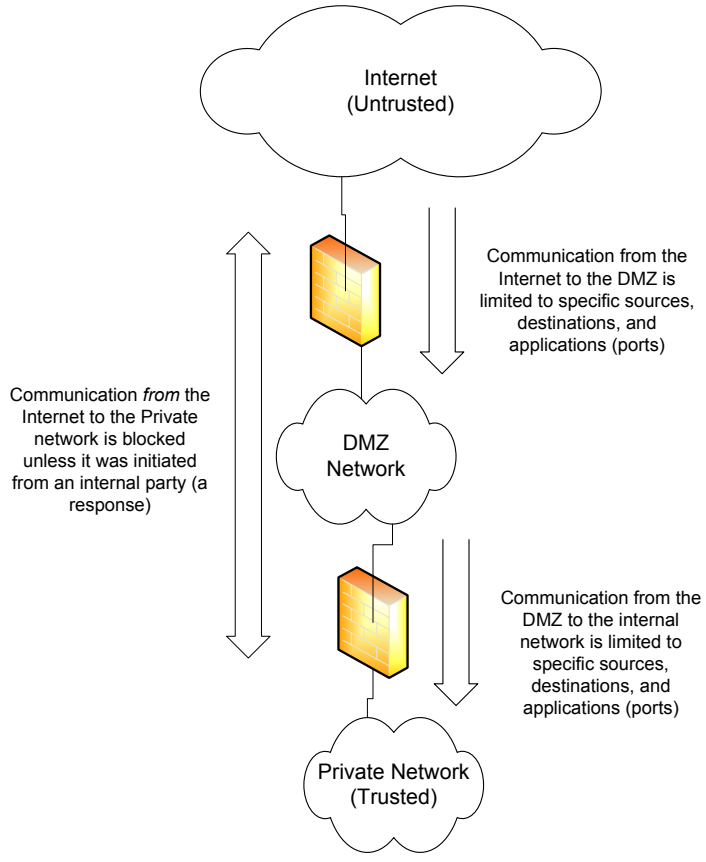
Most firewalls operate by *packet inspection*. For every inbound packet, a firewall makes a decision after examining the packet header whether to forward the packet or discard it. Firewalls make their decisions based on source address (IP and port) and destination address (IP and port). There is a one-to-one correlation between a port and a protocol, so firewalls can allow or block communications between any two IP addresses for any particular protocol. As an example, a firewall could allow any Internet computer to access an internal web server using the http

protocol (port 80), but prevent all Internet users from accessing the same web server using any other protocol (port). These address and port combinations are the configuration parameters for configuring a firewall to allow specific kinds of traffic like VoIP.

In addition to following the address/port rules described above, many firewalls also block some Internet threats such as well-known denial of service attacks. If not properly configured, some firewalls may interpret VoIP traffic as a denial of service attack.

Most firewalls also perform Network Address Translation, or NAT, which presents an additional set of problems for VoIP communications. Refer to the Network Address Translation section for more information.

In many organizations, firewalls provide a special purpose subnet known as a DMZ (DeMilitarized Zone). A DMZ is a network specifically designed for housing Internet servers, and it is logically *between* the internal network and the Internet. There is firewall functionality between the Internet and the DMZ, and again between the DMZ and the internal network. DMZs are the standard location for servers that must be accessible from the Internet, including VoIP servers.



Measurement

The best way to measure a firewall is to test its configuration by attempting the allowed and denied communications. There are numerous port-scan utilities that can help to identify

configuration weaknesses. You can also use a network analyzer to capture the packets on the inside of the firewall to see if undesirable packets are penetrating the firewall.

Tuning

Firewalls are customizable based on a particular enterprise's network security needs. Usually the default configuration is to allow any communication initiated from the trusted internal network and block everything else. To allow VoIP, the firewall can be configured to allow specific protocols (ports) to specific devices (IP addresses). Although firewalls typically allow filtering based on source IP addresses, this is not considered secure because source addresses are easily falsified (spoofed).

The specific detail about the ports and protocols required for VoIP is product and implementation dependent. Depending on the vendor's requirements, this may be as few as one port ranging up to hundreds of ports. Refer to product documentation for more additional information.

Effects

Firewalls are a major component in protecting a network from security threats on the Internet. However, the proper configuration of a firewall is critical. If a firewall is configured to be too strict in its filtering, it can cause calls to fail (call control) or no-audio problems. Depending on the configuration, the problems may be repeatable or intermittent. If a firewall is configured to be too lenient in its filtering, it can leave DMZ or internal devices vulnerable to Internet security threats like hackers.

Note: *Internet security is critical to every organization. Security breaches can result in reduced availability, lost productivity, and compromised information. Making changes to the security configuration should be made with great care by qualified personnel only.*

A firewall may add to the network latency (delay) and network infrastructure (cost). However, a well provisioned and maintained firewall can reduce the total cost of ownership of a network by protecting the network from external threats.

Frame Relay Networking

Definition

A telecommunication service designed for cost-efficient data transmission (often less expensive than point-to-point transmission, especially over long distances). Frame relay networks are digital networks in which different logical connections share the same physical path and some logical connections are given higher bandwidths than others.

Frame relay puts data in a variable-size unit called a frame and leaves any necessary error correction (retransmission of data) up to the endpoints, which speeds up overall data transmission. Frame relay is a fast packet (switching) technology that transmits data without any error checking correction along the route. Assurance that the packet arrived without error is the responsibility of the receiver. Fast packet transmission is possible because of the extremely low incidence of error or data loss. However, when an error is detected in a frame anywhere in the network, it is simply "dropped" (thrown away). The endpoints are responsible for detecting and retransmitting dropped frames. This behavior requires special consideration when implementing VoIP due to the connectionless protocol being used.

Measurement

Note: The following is not intended to be a complete explanation on measuring and calculating frame relay behavior. For additional details visit http://www.mplsforum.org/frame/.

Frame relay circuits are measured by port speed (circuit speed) and committed information rate (CIR). The User to Network Interface (UNI) port is commonly provisioned in 64Kbps bandwidth increments and determines how fast data will traverse the network, and CIR is defined as how much bandwidth the carrier guarantees the circuit lessee (subscriber).

There is usually disparity between the port speed and the CIR that is identified in the contract between the carrier and the lessee (subscriber). The common difference is that a carrier guarantees the CIR's availability to the subscriber. However, the port speed and CIR are tailored to the specific requirements of each individual customer. It is rare to find two frame relay networks configured exactly the same.

Traditionally, frame relay is used to transmit data between geographically disperse local area networks (LANs). By provisioning a circuit to have a greater port speed than CIR, this allows the data subscriber to utilize the entire port (exceed CIR or burst) with the understanding that every packet that exceeds the CIR will be tagged as discard eligible (DE). When congestion is present on the frame relay circuit, the DE-tagged frames are dropped to reduce congestion. Furthermore, with regards to congestion, frame relay uses forward explicit congestion notifications (FECNs) and backward explicit congestion notifications (BECNs) to control the flow of data between edge devices. The FECN and BECN frames are sent between the edge devices to alert the edge device of which direction congestion is detected, thus allowing the edge device to control its data flow. Be aware that CIR, FECN, & BECN are all configured in

software and the traffic shaping (configuration of data flow) associated with each is generally tailored on a per site basis.

Tuning

Tuning a frame relay data flow is known as traffic shaping. Consider the following when traffic shaping Voice over IP (VoIP):

- Provision frame relay circuits with CIR for all VoIP resources to operate simultaneously. In other words, don't burst voice packets as this will result in dropped packets.
- Configure traffic shaping to utilize the FECN and BECN frames.
- If the frame relay port speed is lower than 768Kbps, be sure to configure fragment and interleaving of frames to best utilize the low-speed link.
- Prioritize audio and voice signaling traffic.

Effects

Properly provisioned frame relay circuits can support high quality VoIP. Inadequate CIR directly impacts voice quality.

Internet Precautions

Definition

The Internet, which evolved from the ARPANET of the late 60's and early 70's, is a vast, worldwide collection of independent interconnected networks that all use the TCP/IP protocols to facilitate data transmission and exchange.

Measurement

You can measure upload and download speeds between various points on the Internet, but these measurements often vary due to changes in traffic and routing.

Tuning

Although you can adjust your bandwidth connection (pipeline) to the Internet, you cannot control the data transmission capacity or speed of the Internet itself. Any QOS settings deployed on a controlled network have no effect once data moves on to the public Internet.

Effects

Due to the volatility in delay, jitter, and packet loss, sending VoIP traffic over the Internet will result in intermittent poor voice quality, dropped calls, and equipment resets. No delivery guarantee can be made for VoIP traffic over the Internet.

Also, as a public network, the Internet is subject to a wide variety of security and privacy threats.

Jitter

Definition

Network delay is the average time it takes packets to travel through a network from one host to another. The speed of physical connections, routers, and switches are the major contributors to network delay. Jitter, in VoIP terms, is the short-term variation of network delay.

IP packets compete for bandwidth as they travel through a network. At each router or switch hop, there is a chance that a packet will wait in a queue for the transmission of other packets. Jitter is very low when the queuing delay is consistent for all packets. However, queuing delays can vary greatly as congestion increases which typically results in increased jitter. Normal data traffic has a tendency to be bursty which can cause jitter by changing queue delays.

Jitter is a very important topic for VoIP. Jitter is one of the three main network qualifiers (the other two being latency and packet loss). Excessive jitter can cause packets in real-time audio streams to arrive late, which means that the receiver cannot play them. Late packets have the same undesirable effect on audio quality as lost packets, which can cause choppy or broken audio.

Applications that rely on real-time audio streams use jitter buffers to compensate for jitter. Jitter buffers attempt to absorb jitter by introducing a small buffer at the receiver to wait for delayed packets. Packets with less jitter than the jitter buffer size will leave the jitter buffer at a consistent rate, usually to be played by the receiver. Other packets with too much jitter for the jitter buffer are identified as late packets. The receiver usually discards late packets. Increasing the jitter buffer size allows it to handle audio streams with more jitter at the expense of increasing the overall delay of the audio stream.

Please see the document entitled, "Inter-Tel's VoIP Data Network Requirements" for more details about jitter requirements for VoIP.

Measurement

Applications measure jitter by sending a consistent stream of packets across a network and calculating the variance in the arrival times. For example, one side sends packets at 30 ms intervals with relative timestamps +30, +60 and +90. The receiver sees the three packets arrive with relative timestamps of +10, +60 and +100. The receiver would estimate the average jitter to be about 10 ms ((|10 - 30| + |60 - 60| + |100 - 90|) / 3) and would need a jitter buffer of at least 20 ms to correctly receive all of the packets.

In practice, it is not feasible for a receiver to store all of the inter-packet intervals. Receivers typically use a weighted average to estimate the average jitter of the most recent packets. Using this method, devices such as the Inter-Tel IPRC and many IP phones display the estimated average jitter on diagnostics screens.

Many tools are available to simulate audio streams and measure jitter. The Inter-Tel Network Qualifier measures jitter with this method and simulates a jitter buffer to estimate late packets. Using the common ping utility, a user can examine the variance in the ping response times for a rough estimate of the current jitter. Unfortunately, using the ping utility to estimate jitter has limited accuracy and should only be used to test for the presence of jitter. Use a tool designed to measure audio stream jitter for a more accurate measurement. Networks that implement QoS to give high priority to audio packets may skew or invalidate the measurements of some tools.

Like packet loss, the overall jitter of a network can be difficult to measure because it may greatly increase for very short periods while the network is congested. It is best to run a tool like the Inter-Tel Network Qualifier for many hours during peak usage and examine the output for spikes in jitter. These are the times that jitter will have the greatest effect on audio quality.

Tuning

The most effective ways to reduce jitter are to reduce congestion and implement QoS. Reducing congestion may require increased bandwidth, faster routers, or replacing hubs with switches. There are many methods of implementing QoS. Some of the more common methods are giving priority to audio packets with a certain Differentiated Services value, port ranges, or IP addresses.

Most applications with a jitter buffer provide some configurable settings. The jitter buffer size is a trade off between audio delay and the buffer's ability to compensate for jitter. Static jitter buffers attempt to maintain a fixed buffer size throughout a call. The buffer size is usually configurable before the call begins. It is good practice to maintain a jitter buffer size of at least twice the average jitter. Dynamic (adaptive) jitter buffers attempt to measure the jitter themselves and adjust the buffer size to compensate for it. Dynamic jitter buffers sometimes have parameters that indicate their minimum and maximum values and how quickly the buffer size can change.

Effects

Excessive jitter typically results in poor/choppy audio problems. Often, jitter is mistaken for packet loss when poor audio occurs. However, jitter does not affect reliable connections such as call control and typical data connection as adversely as packet loss.

Increasing jitter buffer sizes can make audio streams more robust with jitter at the expense of audio delay.

IP streaming audio settings which increase bandwidth consumption can cause network congestion and increased jitter. These settings include vocoder and packet size. Modifying these settings can affect audio quality and delay.

Line Echo

Definition

Echo is when you hear your own voice delayed while you are talking. Echo is typically caused by electronic audio reflections at 2-wire-audio-to-4-wire-audio interfaces (called ***"hybrids"***) coupled with "long" delays in the round-trip audio path. This type of echo is called line echo. Another type of echo is acoustic echo and this is usually caused by speakerphones. Line echo is the most common form of echo and hence, we focus on that type of echo here.

A 2-wire-audio device is one in which both the transmit audio and the receive audio travel over the same 2 wires. A 4-wire audio device is one in which the transmit audio travels over a separate pair of wires from the receive audio. The most common example of a 2-wire interface is a standard analog telephone. Another example is a 2-wire analog trunk from the telco such as a loop-start or ground-start trunk.

- All 2-wire-audio trunks and devices *must be analog*.
- All digital trunks and devices *must be 4-wire audio*.

The most common example of where hybrids (and reflections) are found is where a standard analog telephone device connects to a PBX or the telco. Hence, hybrids are found everywhere in the PSTN. This also means that audio reflections are ubiquitous throughout the PSTN.

The reflections are not heard as "echo" unless the round trip delay is "long" enough. To summarize: Echo = reflection + "long" delay [you must have both ingredients to have "echo"]

In VoIP networks, the "long" delays can come from two sources:

- The encoding/compression/packetization and decoding/decompression/depacketization
- Delays across a WAN

Round-trip delays (between the talker and the hybrid) greater than 50 milliseconds will be perceived by the user as "echo." Delays less than 5 milliseconds will not be heard as echo but will be heard as "sidetone." Round-trip delays in-between 5 and 50 milliseconds are a "gray area" and are "subjective" in how they sound.

There are two types of echo:

- "Near-end" echo – echo caused by the near-end equipment
- "Far-end" echo – echo caused by the far-end equipment

Applicable standards that address echo include, but are not limited to the following:

- ITU-T G.114 – One way transmission time
- ITU-T G.164 – Echo Suppressors

- ITU-T G.165 – Echo Cancellers
- TIA/EIA/TSB116 - Voice Quality Recommendations for IP Telephony

For a more in-depth discussion of understanding and troubleshooting echo, see the document entitled "Echo Troubleshooting Guide" found on *the edGe*.

Measurement

Although it's certainly possible to measure the level and delay associated with a particular echo, it really isn't worth doing from a troubleshooting standpoint. Echo is an impairment that is pretty much an "all or nothing" proposition. If the user hears any echo at all, no matter how long the delay or how loud the echo, the effect can be annoying and bothersome.

The proper procedure for eliminating echo is as follows:

1. Create a detailed network diagram
2. Show all known possible reflections (i.e., all hybrids)
3. Show all significant delays (delays exceeding 10 milliseconds)
4. Show all devices experiencing echo
5. Trace the call flow on your diagram for calls experiencing echo
6. Identify sources of echo, where echo = reflections + significant delays
7. Determine whether the echo is "near-end" or "far-end" echo

Once you have identified the sources of echo using your network diagram and call flow, there are three ways to eliminate "echo":

1. Eliminate 2-wire-audio trunks and devices wherever possible
2. Eliminate "long" delays (practically impossible on a VoIP network)
3. Apply "echo cancellers" to all hybrids to stop the echo

If you are not sure whether echo cancellation is applied at a hybrid, contact the equipment manufacturer to inquire about their echo cancellation used for the suspect equipment.

In the case of "near-end" echo, you are responsible for eliminating your equipment's echo. However, in the case of "far-end" echo, you may need to purchase a separate piece of echo-cancellation equipment specifically designed to eliminate "far-end" echo. Far-end echo will require an echo canceller with a tail length of 128 milliseconds.

Tuning

Although there are a few things you can "tune" to reduce the reflections and the delays that contribute to echo, in general, you cannot "tune" away echo. Here are the areas you can adjust:

Hybrid Balance settings – at any 2-wire-audio interface, such as a loop-start analog trunk card, if there is a configuration adjustment for the Hybrid Balance, you may be able to

reduce the level of the echo by selecting a better hybrid balance setting, but you probably won't be able to eliminate it altogether.

Delays – there are a few things you can do to minimize delays, but for VoIP networks, you probably will not be able to minimize the delay enough to eliminate echo. This is because the minimal packetization delay will be enough to cause echo. But shorter delays are definitely desirable and should be pursued. For more details about minimizing delays, see the section entitled ***"Audio Delays."***

Echo Canceller settings – in some cases, you may be able to make slight adjustments to your echo canceller (and echo suppressor) settings in your equipment. Check with your equipment manufacturer's technical support staff before attempting any changes to the echo canceller settings.

Unlike other VoIP audio impairments, such as choppy audio, echo has its roots in the analog domain and not the data network. As such, the solution to eliminating echo also has nothing to do with the data network. That is, adjusting bandwidth, packet loss, QoS, etc. will have no effect on eliminating echo. The only data network adjustment that will have any impact on echo is the topology itself. Since echo is a result of delay, minimizing delay in your VoIP network can help reduce the magnitude of the echo's delay. However, even in the scenario of a switched LAN with negligible delay between devices, the packetization delay is enough to cause some echo if reflections occur in the audio path. Hence, the best solutions to eliminating echo are:

1. Eliminate hybrids wherever possible (i.e., eliminate all 2-wire device)
2. Use echo cancellers at all hybrids

Effects

Adjusting hybrid balance settings can make line echo louder or softer. In some cases, it may even eliminate echo. However, in most cases, adjusting hybrid balance settings will, at best, make the echo softer, but not eliminate it.

Adjusting delays (through frames per packet, vocoder selection, jitter buffer size, and network topology) can make the echo have shorter or longer delays. Shorter delays are more desirable. However, it will probably be impossible to make the end-to-end delay short enough to eliminate echo altogether.

It may appear that an echo canceller works fine on some calls but not on other calls. This may be because the echo canceller's tail length is not long enough to cancel echo for calls involving far-end echo. Because each call destination can be different, every call involving echo may have different echo. If the echo canceller's tail length is adequate to cancel the echo, it should do so. In general, echo cancellers do not requiring "tuning" and should be pretty much transparent to the user. However, if an echo canceller appears to not be working all of the time, it may be due to the echo canceller's tail length. Contact the manufacturer of the echo canceller for additional support should the echo canceller fail to perform as expected.

Network Address Translation (NAT)

Definition

Network Address Translation (NAT) is an Internet standard (RFC3022) that enables a local-area network (LAN) to use one set of IP addresses for internal traffic and a second set of addresses for external traffic. Typically, the NAT function is provided by a router or firewall. (Occasionally, the device that performs Network Address Translation is referred to as a NAT). NAT can present significant problems for VoIP communications because VoIP call control packets usually contain IP addresses in their *payloads* (vs. IP addresses in packet headers).

The main purpose of NAT is to allow an organization to use a virtually unlimited pool of (private) IP addresses that is separate from (public) Internet IP addresses. This alleviates a shortage of public IP addresses. Although NAT can provide a limited amount of security, it is rarely used on its own for security purposes.

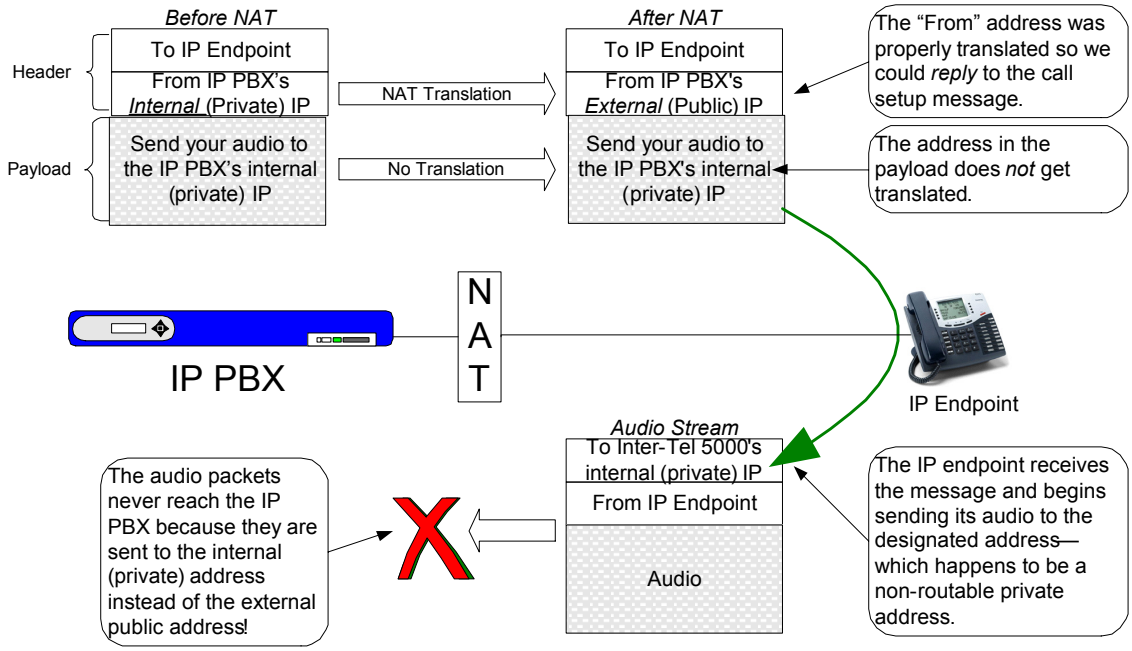
To facilitate NAT, the Internet Assigned Numbers Authority (IANA) has designated certain IP addresses to be *private*. This designation means that these IP addresses are not valid (routable) on the public Internet. This allows multiple organizations to safely utilize these addresses within their own networks. The designated private addresses are:

10.0.0.0 - 10.255.255.255	(One Class A Subnet)
172.16.0.0 - 172.31.255.255	(16 Class B Subnets)
192.168.0.0 - 192.168.255.255	(256 Class C Subnets)

NAT operates by dynamically associating each internal private IP address with an external public IP address (and port) as a packet goes out to the Internet. The NAT device (router or firewall) keeps track of the association between internal and external addresses and re-writes the IP packet header addresses as necessary. The association between internal and external IP addresses is typically short-lived based on activity.

Although NAT has long been widely used throughout the Internet, some protocols like VoIP do not work well with NAT. As described above, NAT translates the IP addresses in only the *headers* of the IP packets. The root of the problem is that some protocols carry IP addresses in the *payload* of the IP packets. The result is that private IP addresses are sometimes communicated in the payload out to the public Internet. By definition, these private IP addresses are not accessible.

To illustrate the issue caused by NAT, consider the following example of an IP endpoint receiving an inbound call. When the endpoint answers the call, the IP PBX sends a packet to an IP endpoint indicating *where* (IP address) to send audio.



As a result of these problems, some Internet applications (e.g. IP telephony) forbid the use of NAT. Some vendors of firewalls offer capability to "fix" the NAT problems for specific protocols (e.g. SIP) by interpreting the protocol and translating the IP addresses where appropriate throughout the payload.

It is possible to configure persistent or *static* NAT assignments in which a specific outside address is associated with a specific inside address. Although static NAT essentially allows an inside device to be accessible from the outside, it is still NAT and therefore problematic for VoIP protocols in which IP addresses are carried in the payload of the IP packets.

Because NAT takes place where a private network connects to a public network, there are often two NAT operations taking place—one at each end of the communication. The NAT operations are the same at the two ends, but sometimes the impact on network protocols is different depending on which end embeds IP addresses in the payload of the packets. These are described as near-end NAT and far-end NAT depending on your perspective.

There is a variation of NAT called PAT, Port Address Translation, in which the trusted IP address and port are translated dynamically to a single an outside IP address and (different) port. In this way, PAT leverages fewer public IP addresses than NAT. Devices that perform both NAT and PAT, sometimes referred to as *overloading*, will typically use exclusively NAT until they run out of external addresses. For VoIP applications, NAT and PAT present the same problems.

Measurement

Most devices that perform NAT allow you to examine the NAT (and/or PAT) associations. In the case of NAT, you can see the association between internal IP address and external IP address. For PAT, you can see the association between internal IP address *and port* and external IP address *and port*.

Tuning

NAT and its impacts on VoIP communications need to be considered in network topology design. Whenever practical, VoIP devices should be placed in a network so that NAT traversal is minimized.

The lifespan of the dynamic NAT association is typically configurable. In most cases, the lifespan is not a major consideration for VoIP, and the default configurations are appropriate.

Effects

NAT and PAT minimize the number of public IP addresses required for a private network connected to the Internet. However, the simple process of translating addresses in the packet headers is not adequate for all types of communications—especially VoIP.

Network Topology & Diagramming

Definition

A proper network topology diagram is a detailed diagram that provides a schematic description of the arrangement of a network, including servers, routers, switches, endpoints, etc., and all connecting circuits.

An example of a detailed network diagram is shown on the following page.

Measurement

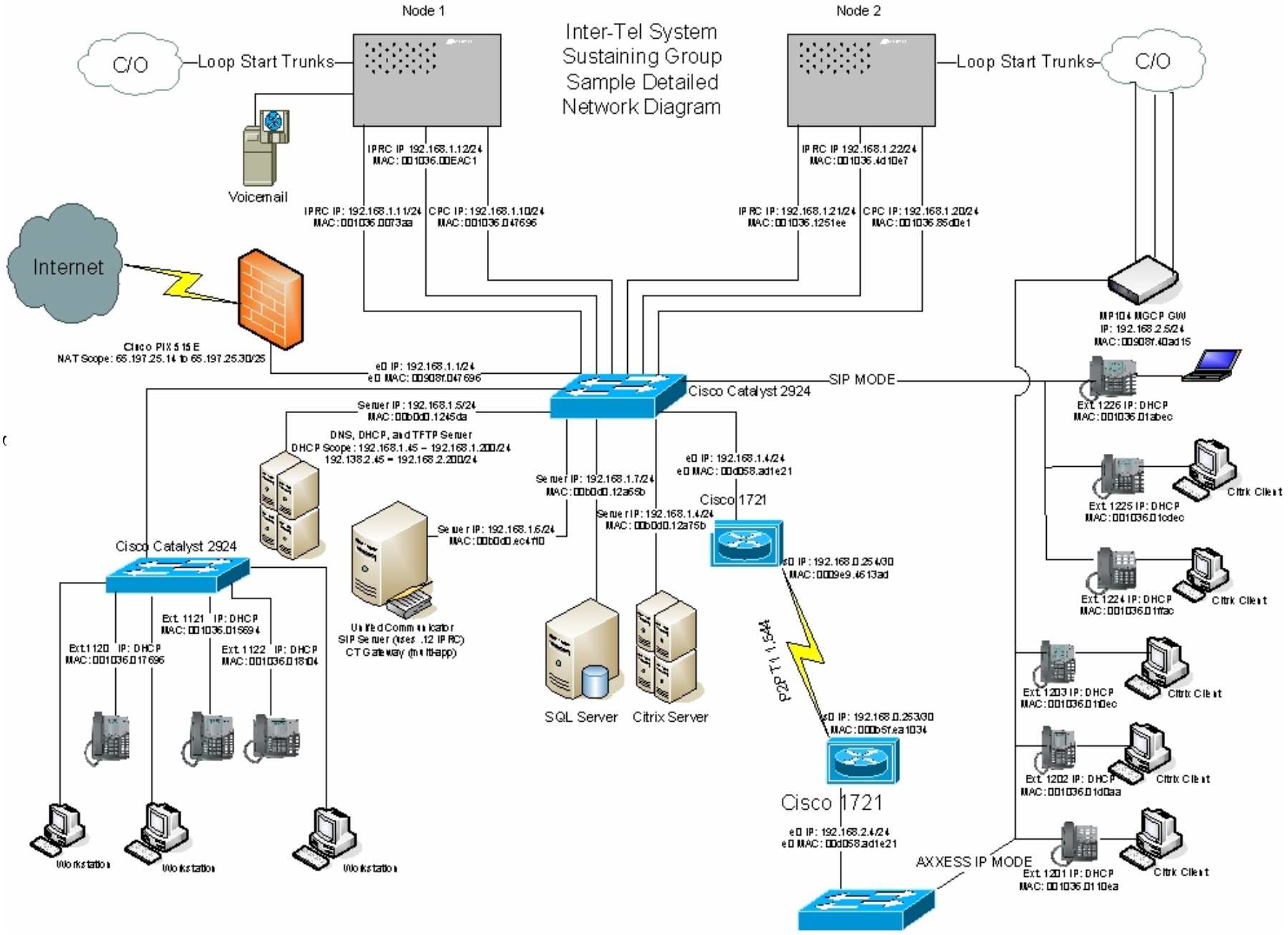
There is really no measurement beyond the extent of detail. A good network diagram includes low-level detail, such as exact number and model of endpoints, subnet information, circuit capacities, static IP addresses, host names, DHCP server address, DNS server address, outbound/inbound firewall ports, etc.

Tuning

Unless the network is being designed from the ground up and is a new implementation, tuning consists of adding details as they are discovered.

Effects

A detailed network diagram results in improved planning and troubleshooting capabilities.


Inter-Tel System
Sustaining Group
Sample Detailed
Network Diagram
Node 1
Node 2
C/O
Loop Start Trunks
Loop Start Trunks
C/O
Voicemail
IPRC IP 192.168.1.12/24
MAC: 001036.00EAC1
IPRC IP 192.168.1.22/24
MAC: 001036.4d1De7
IPRC IP: 192.168.1.11/24
MAC: 001036.0073aa
CPC IP: 192.168.1.10/24
MAC: 001036.047696
IPRC IP: 192.168.1.21/24
MAC: 001036.1251ee
CPC IP: 192.168.1.20/24
MAC: 001036.85d0e1
Internet
Cisco PIX 515E
NAT Scope: 65.197.25.14 to 65.197.25.30/25
MP104 MGCP GW
IP: 192.168.2.5/24
MAC: 00907f.40ad15
e0 IP: 192.168.1.1/24
e0 MAC: 00907f.047696
Cisco Catalyst 2924
SIP MODE
Server IP: 192.168.1.5/24
MAC: 00b0d0.1245da
DNS, DHCP, and TFTP Server
DHCP Scope: 192.168.1.45 – 192.168.1.200/24
192.138.2.45 – 192.168.2.200/24
Ext 1226 IP: DHCP
MAC: 001036.01abec
e0 IP: 192.168.1.4/24
e0 MAC: 00d058.ad1e21
Server IP: 192.168.1.7/24
MAC: 00b0d0.12a65b
Server IP: 192.168.1.4/24
MAC: 00b0d0.12a75b
Cisco 1721
Ext 1225 IP: DHCP
MAC: 001036.01cdec
Citrix Client
Server IP: 192.168.1.6/24
MAC: 00b0d0.ec4f10
Cisco Catalyst 2924
e0 IP: 192.168.0.254/30
MAC: 0009e9.4613ad
Ext 1224 IP: DHCP
MAC: 001036.01ffac
Citrix Client
Ext 1121 IP: DHCP
MAC: 001036.015694
Ext 1120 IP: DHCP
MAC: 001036.017695
Ext 1122 IP: DHCP
MAC: 001036.01810d
Unified Communicator
SIP Server (uses .12 IPRC)
CT Gateway (multiapp)
SQL Server
Citrix Server
P2P T1 (54)
Ext 1203 IP: DHCP
MAC: 001036.01dec
Citrix Client
e0 IP: 192.168.0.253/30
MAC: 0006.5fea1034
Ext 1202 IP: DHCP
MAC: 001036.01d0aa
Citrix Client
Cisco 1721
e0 IP: 192.168.2.4/24
e0 MAC: 00d058.ad1e21
AXXESS IP MODE
Ext 1201 IP: DHCP
MAC: 001036.01d0ea
Citrix Client
Workstation
Workstation
Workstation

Packet Loss

Definition

IP networks transport packets between hosts through a series of physical connections, switches, and routers. Packet loss occurs when packets fail to arrive intact at the destination. It is normal for most IP networks to have a small amount of packet loss that higher level protocols and applications must compensate for.

The method that applications use to handle packet loss usually depends on the type of data being transported. Most applications require reliable data transmission and use protocols such as TCP or a reliable form of UDP to do error detection and data retransmission. Examples of these applications are web browsers, file transfer utilities, and IP phone call control. Some applications do not require completely reliable data but are time dependent where retransmission is not an option. Often, the application performance degrades with data reliability. Streaming media such as VoIP audio is a common example.

In the case of VoIP, packet loss is a very important topic. Packet loss is one of the three main network qualifiers for VoIP (the other two being jitter and latency). Packet loss has a direct impact on the quality of voice. The majority of VoIP audio problems reported by the customer will usually be tied directly to excessive packet loss. For this reason, it is critical that the troubleshooter understand the "symptoms" of packet loss, how to identify the sources of packet loss, and the appropriate solutions for packet loss.

Please see the document entitled, "Inter-Tel's VoIP Data Network Requirements" for more details about packet loss requirements for VoIP.

Measurement

Applications use a variety of methods for detecting packet loss. Some of the most common involve including sequence numbers in every packet sent or requiring the receiver to acknowledge each packet. Some applications, such as IP endpoint audio streams, display the packet loss that they detect in their data. Other applications transmit simulated data with unreliable protocols for the sole purpose of measuring how much is lost. Network sniffers may be able to measure packet loss by examining the packets of other applications.

Many tools are available that simulate data and measure packet loss. The Inter-Tel Network Qualifier simulates streaming audio data between two hosts and displays the percentage of received packets. The common ping utility can estimate the round trip packet loss by echoing ICMP packets off a remote host. Unfortunately, the ping utility does not indicate the direction, request, or response of the loss occurred. Networks that implement QoS and give low priority to ICMP packets may make measurements using the ping utility less accurate.

Measuring packet loss can sometimes be misleading. It is important to understand that a packet loss of 1% as measured over 1 month (as some providers do) is not a guarantee of good voice

quality. This is because packet loss can be burst packet loss. Any short term or burst packet loss that exceeds the VoIP requirements for good quality can result in poor audio quality. Hence, it is very important to note whether the packet loss being measured is burst or uniformly distributed packet loss.

Tuning

To reduce packet loss, determine where the majority of loss occurs by examining each hop in the route. The common traceroute utility is sometimes useful for this task. Next determine if the loss is caused by physical link errors or network congestion. Typically, physical link errors cause random packet loss while network congestion causes burst packet loss. Physical links that generate excessive errors require servicing. An administrator can usually improve loss due to congestion by increasing bandwidth, replacing hubs with switches, and/or implementing QoS methods depending on where the congestion occurs.

Some applications are configurable to help manage packet loss. Adjusting the packet size and vocoder on IP endpoints directly affects bandwidth utilization which can impact congestion and packet loss. Some reliable connections such as IP phone call control have configurable timeout and retransmit settings. These settings can allow additional time to retransmit lost data but can sometimes make applications less responsive.

Excessive jitter can cause data loss in unreliable connections such as VoIP audio streams. Data loss due to jitter differs from packet loss because the packets arrive at the destination but the packets are too late to be useable. However, the effect on audio quality due to excessive jitter is very similar to the effect of packet loss. Please refer to the jitter section for more information.

Effects

Poor/choppy audio, lost connections and unresponsive applications are problems typically associated with packet loss.

IP streaming audio settings which increase bandwidth can cause network congestion and increased packet loss. These settings include vocoder and packet size. Modifying these settings can affect audio quality and delay.

Increasing connection timeouts and retransmission limits can make reliable connections more robust with packet loss but increase the detection and reconnection time of truly lost connections.

Some applications, T.38 fax for example, have an option to send redundant data to help hide packet loss. While this is helpful for some causes of packet loss such as physical link errors, it greatly increases bandwidth and congestion.

Phone System Security

Definition

Unauthorized access to the phone system may allow the following:

- Unauthorized changes to DB Programming
- Toll Fraud
- Hacking/crashing the system
- Denial of Service (DOS) attacks

Measurement

Appropriately protecting the data network and using protected passwords will minimize the exposure of the phone system. Thus, the greatest concern to the phone system is Denial of Service (DOS) attacks. To execute a DOS attack, malicious users only require being able to send network packets to the target device.

Inter-Tel systems, like all data systems, are susceptible to overload from a malicious source. This can happen by sending too many packets to the system and overwhelming the system. If this were to happen, the IP stations assigned to the system will be out of commission until the system can be restarted, restored, or protected. Additionally, communication to other networked devices may be interrupted. This may have an effect on centralized features, such as networked voicemail.

Tuning

Change all passwords regularly to unique non-default values that follow generally acceptable password complexity rules (i.e. alphanumeric values that include upper and lower cases and are not dictionary terms).

Manage user accounts by mandating that all users change passwords regularly, prevent access to accounts of ex-employees, etc.

Effects

Note: Although no telecommunications system or data network is entirely secure, as long as the appropriate security measures are put in place and properly maintained by both the customer and the installing company (including the proper implementation of user/administrative accounts, passwords, firewalls, NAT, virus protection, security updates, etc. and the proper maintenance of access points/programs and their respective accounts/passwords), the Axxess system architecture and its associated server-based applications are substantially secure against unauthorized access to the customer's data network via the telecommunications system.

Power over Ethernet (802.3af)

Definition

Power over Ethernet generally describes any system that transmits electrical power and data to devices over standard twisted-pair wire in an Ethernet network. This eliminates the need to provide separate network and power connectivity to low power devices, such as IP telephones.

The IEEE 802.3af specification standardizes PoE. It specifies a power supply of 48 volts DC at a maximum current of 350 mA. This allows devices to draw about 15.4 watts. IEEE 802.3af compliant power supply will not send power to non-compliant devices to prevent damage to devices that aren't designed for PoE.

The IEEE 802.3af specification itself can be found at:
http://standards.ieee.org/getieee802/download/802.3af-2003.pdf.

Measurement

Make sure that the Ethernet switch used is 802.3af compliant.

Tuning

Single port PoE devices (i.e. "bricks") generally don't follow the IEEE 802.3af standards and can damage attached equipment if said equipment cannot handle the voltage.

Effects

PoE offers flexibility for users and administrators by removing the need for client-side power devices and allowing small devices to be powered via the same cable as its Ethernet connection. If the PoE switch is backed up by an uninterruptible power supply, attached PoE devices, including VoIP endpoints, can maintain their operation in the case of power outage. (Note that the rest of the core network and PBX infrastructure requires the same battery-backing to ensure the availability of the VoIP telephony system is such an event.)

The use of PoE power sources that are not IEEE 802.3af compliant can damage devices that do not support or handle receiving power via its Ethernet connection.

Switches and Hubs

Definition

Switches and hubs are both devices that provide connections to multiple Ethernet devices like PCs, servers, and IP endpoints. Although switches and hubs share the same external appearance (multiple RJ45 jacks), they operate considerably differently internally. The choice of which device to use can significantly affect VoIP audio quality.

A *hub* is a relatively simple device that simply retransmits every *bit* it receives on *any* of its ports to *all* of its other ports. Although hubs are extremely fast (one-bit delay), all connected devices share bandwidth and can potentially cause collisions when attempting to transmit frames. In other words, all devices connected to a hub share the same *collision domain*.

A *switch* is a more intelligent device that attempts to perform a "routing" function at the Ethernet level so that data goes only where it is needed (instead of everywhere). Unlike hubs which forward *bits*, switches operate at the *frame* level. By watching *source* MAC addresses in frames, switches can keep track of which devices are connected to which ports. Instead of blindly forwarding every frame, switches forward frames to only the appropriate port(s). Because each Ethernet port on a switch sees traffic *to* the connected device or *from* the connected device, the devices on different ports do not compete for bandwidth. In this way, a switch divides a network into multiple collision domains. This increases the likelihood of VoIP traffic getting through to its destination without delay.

Switches are much faster devices than a similar device called a *bridge* because switches can usually begin forwarding the frame as soon as they've seen the destination MAC address at the beginning of the frame (i.e. a six-byte delay vs. one frame delay). The ability to begin transmitting before receiving the entire frame is referred to as *cut-through* switching. Although each port essentially connects to a separate Ethernet network (with just two devices), it is possible that the *other* device is transmitting which would cause the switch to wait before transmitting. Cut-through switching can happen only if the destination port is able to transmit. In addition, many switches can support different ports providing different bandwidths (e.g. 10Mbps, 100Mbps), so it's simply not possible to transmit the frame as fast as it is being received. When cut-through switching is not possible for a particular frame, the switch falls back to slower *store-and-forward* functionality in which the entire frame must be received and queued before transmission can begin (i.e. a one-frame delay). Frames that are "stored-and-forwarded" are subject to queuing delays. This can increase delay and/or jitter and therefore may reduce VoIP audio quality. Although queuing delays in an Ethernet switch are rarely problematic, some switches support IEEE 802.1q QoS protocol that allows prioritization of frames when there is queuing.

When there are only two devices on a network, the devices no longer need to take turns transmitting. The first device's transmit channel can be the second device's receive channel and vice-versa. In this scenario, the media access control is not necessary, and both devices can

transmit at the same time, also known as *full-duplex*. Since full-duplex Ethernet communications utilize a different protocol (no media access control), full-duplex must be negotiated between the switch and the device or manually configured in both places. Since full-duplex virtually eliminates queuing delays at the Ethernet level, it is the preferred method for VoIP traffic.

Although switches segregate traffic into multiple collision domains, they are not able to filter *broadcast* frames because they are addressed to *all* devices. Switches must therefore forward broadcast frames to all ports. The set of devices that share broadcasts is known as a *broadcast domain*. In some networks, excessive broadcasts can impact device performance and consume LAN bandwidth. To subdivide broadcast domains, you must separate the network into multiple LANs using a *router*. Many business-class switches support *VLAN* capability which allows a set of switch ports to be treated as a virtual LAN. VLANs, like physical LANs, communicate with each other through a router.

It should also be pointed out that there is often a difference between business-class Ethernet switches and consumer-class Ethernet switches. Consumer switches are less likely to support several features including bandwidth conversion, full-duplex, manageability, QoS, VLANs, bandwidth aggregation, etc.

For nearly all applications, including VoIP, switches are strongly favored vs. hubs for the following reasons:

- Switches increase Ethernet throughput by subdividing the collision domain.
- Switches can provide full-duplex communications allowing devices to transmit at the same time they are receiving.
- Business-class switches often provide QoS capabilities.
- Business-class switches support usually support multiple bandwidths at the same time.
- Business-class switches are usually manageable which provides greater visibility and control of Ethernet traffic.

Measurement

The communications on an Ethernet network are best measured using a LAN protocol analyzer (e.g. Ethereal, Sniffer, etc.) Protocol analyzers capture *all* frames on the media regardless of MAC address. Note that using a protocol analyzer in a switched environment is more difficult because switches are specifically designed to segregate the traffic according to destination MAC address. It may be necessary to utilize a hub or a switch feature known as *port mirroring* to allow a protocol analyzer to collect the appropriate information.

Most Ethernet switches provide a visual (LED) indication of port bandwidth, full duplex, and network traffic. Although these indicators are simple, they are often examined as the first step troubleshooting.

The important measurements on an Ethernet network include network utilization and collisions. Many Ethernet devices themselves, including switches, provide counters of frames sent and received, collisions, etc.

Tuning

If the measurements indicate LAN congestion, the first step in tuning is usually to identify where hubs are used and replace them with switches. In nearly all applications, switches will perform significantly better than hubs. Although switches are intelligent devices, in most cases they require little or no configuration.

The second step is to identify the connections *between* switches. This will identify potential traffic bottlenecks. As an example, connecting two 100Mbps switches together with a single 100Mbps connection probably is going to cause a bottleneck because all of the devices on each switch will be competing for the 100 Mbps bandwidth to the other switch. Note that some switches, by design, will prevent multiple paths between switches (i.e. loops) through the use of a spanning tree algorithm. Care should be taken to avoid loops because spanning tree algorithms are sometimes disabled, and the traffic resulting from a loop can bring a network down. Many business class switches provide higher-bandwidth "uplink" ports or the ability to aggregate multiple ports together for the interconnections.

In most cases, significant performance improvements can be gained by utilizing full-duplex connections, particularly to server devices (including VoIP servers). If LAN congestion remains a problem, some switches provide the capability to give priority to voice traffic through IEEE 802.1q QoS queue prioritization.

Effects

Switches are intelligent devices that improve network throughput by attempting to restrict Ethernet traffic to only the needed parts of the network. A properly designed Ethernet network utilizing 100Mbps full-duplex switched connections should be capable of providing the bandwidth necessary for VoIP.

TCP/IP Protocols

Definition

A protocol is a set of rules for communications. Internet Protocol (IP) is the protocol for communicating on the Internet. TCP/IP is actually a set of protocols, based on IP, that are commonly supported. One of the main design goals of the TCP/IP protocol was to be very resilient to network outages. The robustness of the protocols has made them the de facto standard for network communications for private networks such as corporate LANs and WANs. Because of its prevalence, IP has also become the de facto standard for carrying voice on a packet switched network.

The TCP/IP protocols were originally designed for *data* communications such as file transfers and e-mail. *Voice* over IP requires some timing-related characteristics not present in the original design. As a result, IP networks adequate for data applications may not be adequate for VoIP without additional controls like QoS.

TCP/IP is based on a layered architecture. Transmission Control Protocol (TCP), User Datagram Protocol (UDP), and Internet Control Message Protocol (ICMP) are all layered, as peers, on top of IP as shown in the following diagram. They are sometimes referred to as "layer 4" or "transport protocols."

<table>
<tr><td>Layer</td><td colspan="5">Applications (e.g. VoIP)</td></tr>
<tr><td>5</td><td>Web, mail, file transfer, etc.</td><td>Call Control</td><td>RTP (voice)</td><td>Ping</td><td>Tracert</td></tr>
<tr><td>4</td><td colspan="2">TCP</td><td>UDP</td><td colspan="2">ICMP</td></tr>
<tr><td>3</td><td colspan="5">Internet Protocol (IP)</td></tr>
<tr><td>2</td><td colspan="5">Data Link Protocols (e.g. Ethernet)</td></tr>
<tr><td>1</td><td colspan="5">(Physical)</td></tr>
</table>

TCP, UDP, and ICMP differ in the services that they provide to the applications that use them:

- TCP provides reliable transmission of information from one point to another. The TCP layer takes care of requesting retransmissions of lost packets as well as re-ordering of packets. In VoIP, TCP is often used to carry the *control* information such as call setup and termination.

- UDP provides a simple, lightweight means for transmitting information from one point to another. The UDP layer does not provide the reliable delivery mechanisms provided by TCP. UDP is used for communications that are more dependent on timing than reliable delivery. In VoIP, RTP (Real-Time Protocol) runs on UDP to carry the *voice* packets because retransmission of lost packets would be too late to be useful and are thus discarded. (In some VoIP protocols like SIP, UDP can be used for the control

information as well as the voice. In these cases, the application, SIP, is responsible for ensuring reliable delivery).

- ICMP provides a simple protocol for network administrative functions such as ping, routing protocols, etc. ICMP is typically not directly used for VoIP, but it is sometimes used to troubleshoot VoIP problems (e.g. ping, traceroute, etc.). Note that routers typically treat the different protocols, TCP, UDP, and ICMP, differently during times of network congestion by selective discarding certain protocols, so troubleshooting with ICMP is inexact.

Measurement

Although many network devices provide some rudimentary form of accounting, the best way to view the use of protocols is through the use of a network protocol analyzer. Two of the most common are Network General's Sniffer ® and the open-source Ethereal. Protocol analyzers collect information at the Ethernet layer and present the information in a human-readable form.

Using a protocol analyzer, it is possible to measure the network utilization relative to capacity. It is also possible to measure the traffic associated with specific calls by filtering on specific IP addresses, protocols, and ports.

Tuning

In most cases, the applications specify the protocols and it is not configurable. The TCP/IP protocol is controlled by routers and firewalls. In order to ensure acceptable voice quality in networks carrying both voice and data, routers and firewalls should be configured to provide QoS for voice.

Effects

The versatility of TCP/IP allows data applications and VoIP to share a common network. However, the real-time requirements of VoIP are usually in conflict with the bursty nature of most data applications, and VoIP voice quality may suffer unless QoS is properly applied.

Vocoders

Definition

In the context of VoIP, the vocoder describes the way the digitized audio information is represented in a voice packet. The two vocoders of interest to us are ITU G.711 (referred to hereafter as G.711) and ITU G.729 and its variants (referred to hereafter as G.729).

G.711 is the "plain vanilla" method used widely for years in digital telephony. It provides toll-quality speech, but requires a high bandwidth. Because it is the "native format" for digital speech, when it is used in VoIP, the transmitter puts the data into packets "as is" with the appropriate headers; the receiver recovers the speech directly from the packet payload.

G.729 is a method of representing the speech that uses less bandwidth. The trade-off of using less bandwidth is that some distortion may be introduced. The end result is near-toll-quality speech. Because it is not the native format, the payload requires additional processing before it is transmitted (converting from G.711 to G.729) and after it is received (converting from G.729 to G.711).

Measurement

The primary measurement for a given vocoder in this context is the required bandwidth. For the ***payload only***, G.711 requires 64 kbps and G.729 requires 8 kbps. When including the header information, the required bandwidth is approximately ***86 kbps total*** for G.711 and approx. ***32 kbps total*** for G.729. Note that the exact bandwidth will depend on the packetization interval used, as well as the network medium (e.g., Ethernet, Frame Relay, PPP, etc.).

Tuning

The largest tuning effect is the selection of the vocoder. G.711 requires more bandwidth, while providing toll-quality speech; G.729 requires less bandwidth, while still providing near-toll-quality speech.

When considering vocoder choices, note that the two ends of the VoIP call negotiate between themselves to find a common vocoder that both ends are capable of using. (It is not required to use the same vocoder for both direction, but that is usually what happens.) As a result, the vocoder used for a given call may be different from the preference that was selected, depending upon the negotiation between the two ends.

Effects

Selection of the vocoder will have the primary influence on the speech quality and bandwidth required. Select **G.711** for **better voice quality**; select **G.729** for a **lower bandwidth requirement**.

Voice Transmission Security

Definition

Unauthorized users should not be allowed to snoop in and record calls.

With the right tools, snooping in on a VoIP call is only slightly more difficult than tapping into an analog call. If an unauthorized user has access to the physical telephone line, a call can be tapped. If an unauthorized user has access to the data network route that the VoIP call traverses, the call can be snooped.

VoIP transmissions that stay within a secure data network are safe from external snooping. VoIP transmissions that traverse the public network have no such protection.

Use of "encrypted" voice packets may provide additional protection and is roughly equivalent to the deployment of secure TDM voice prioritization on digital telephony.

Measurement

N/A

Tuning

Always use switches, not hubs, to minimize the exposure of the VoIP data to other network users.

Note that Inter-Tel VoIP devices currently do not support encrypted VoIP communications.

Effects

If available, encrypted VoIP data adds to the network bandwidth requirements.

Note: Although no telecommunications system or data network is entirely secure, as long as the appropriate security measures are put in place and properly maintained by both the customer and the installing company (including the proper implementation of user/administrative accounts, passwords, firewalls, NAT, virus protection, security updates, etc. and the proper maintenance of access points/programs and their respective accounts/passwords), the Axxess system architecture and its associated server-based applications are substantially secure against unauthorized access to the customer's data network via the telecommunications system.

Part No. 835.2976
Issue 1, March 2005