

Allworx VoIP White Paper

Copyright © 2009 All Rights Reserved – Allworx, a wholly owned subsidiary of PAETEC Holding. No part of this document may be used or reproduced in any manner whatsoever without written permission, including quotations embodied in critical articles and reviews.

Table of Contents

1. Introduction
2. What Does VoIP Really Mean Anyway?
3. Allworx VoIP Features
4. SIP Protocol and VoIP
5. Echo in VoIP Networks
 - 5.1 Where does the echo come from?
 - 5.2 The need for Line Echo Cancellers in VoIP
 - 5.3 If VoIP systems have a LEC and the PSTN has NEC's, why do we still hear echo?
 - 5.4 Should Gateways use both LEC's and NEC's
6. Conclusions
7. Bandwidth Calculations
 - 7.1. Codec support Matrix
 - 7.2. Capacity Planning
 - 7.3. SIP protocol and NAT Firewalls
8. Allworx Solves SIP NAT Problem
9. QoS on the LAN
10. QoS Across a WAN
11. Key System Feature – Allworx BLF Protocol
 - 11.1. ABLF Protocol Background
 - 11.2. ABLF Protocol Operation
 - 11.3. Troubleshooting ABLF
12. Remote Office Phones
13. Zoning Paging
14. Common Problems and Tips
 - 14.1. One Way Audio
 - 14.2. Intermittent Connectivity or devices dropping off the network

- 14.3. Common MAC Addresses on LAN and WAN ports of a router or Firewall
 - 14.4. DTMF Digits not passed during live calls
 - 14.5. Music on hold Sound quality is poor
 - 14.6. Calling Auto Attendant
 - 14.7. Mapping ports through remote phone firewalls
15. Glossary

1) Introduction

The Allworx family of IP based telephone systems and VoIP phones are designed to meet the communications and networking needs of typical small businesses, while also simplifying the setup and maintenance of the voice infrastructure for business owners. The primary mission of the Allworx product line is to deliver to small businesses the most recent round of IT technologies (including VoIP) in “turnkey” solutions that previously were only practical for large enterprises with full-time administrative teams. While Allworx makes voice solutions much easier to use and deploy, having a high-level grasp of the various technologies involved goes a long way toward understanding VoIP systems better and being able to diagnose issues that may occur.

This paper is intended to be a tutorial on several topics as they relate to deploying and maintaining digital phones on a VoIP network. In particular, this document aims to arm administrators, installers, and network planners with information to help them take advantage of the new technologies associated with moving voice traffic over data networks, both on a LAN and across several sites via a WAN.

2 What Does VoIP Really Mean Anyway?

The term “VoIP” is officially an acronym for **V**oice **o**ver **I**nternet **P**rotocol, but is also used to loosely refer to any application where packet-based data networks are used to *packet switch* telephone calls in real-time. This type of telephony contrasts to traditional hard-wired analog telephony that is *circuit switched* known as TDM (Time Division Multiplexing). Migrating to pure VoIP-based telephony technology like Allworx has several advantages, both technical and economic, but also introduces some new complexities that must be managed as part of the data network.

Traditionally, data networks and the Internet in general were developed only as a “best-effort” service. The network was designed to get the data there as fast as possible – and when there were problems, to get as much data there as possible *eventually*. This is a good design characteristic for data, but it has problems with true real-time constraints to support *toll-quality* telephone calls. For telephone audio, not only are bandwidth and throughput important, but packet loss, latency, and jitter performance are also critical factors to good-sounding audio. Therefore, real-time applications like VoIP gave rise to engineering and managing the **Q**uality **o**f **S**ervice (QoS) of data networks.

Designing networks for QoS factors and diagnosing QoS problems are entirely new dimensions in data networks for many people. In VoIP applications, a valid data connection is required to ensure application success *and* a high-quality, maintainable QoS. It is not uncommon for poorly managed or improperly configured data networks to have throughput or packet loss problems that go completely unnoticed until VoIP systems are deployed. Thus, when deploying VoIP systems, it is important to inspect or validate the existing network to make sure it is going to be VoIP-ready from a QoS perspective. This is an especially important consideration when VoIP calls are going to be placed over data connections between physical locations. QoS topics are further explored later in this document, but the main items of interest are network packet latency, jitter, and packet loss rates. Latency is the amount of time that a data packet takes to get from 1 point across a data network. IP packets can take many different routes (known as “hops”). Reassembling the data packets into voice conversation is time sensitive, and when packets are received after a defined threshold has passed, the loss of those packets results in jitter. Both LAN and WAN factors influence latency and jitter.

3 Allworx VoIP Features

The Allworx family of servers is sophisticated VoIP PBX's and gateways designed specifically for use by small businesses. The 6x has a maximum capacity of 60 users, and the 24x is designed with a maximum capacity of 150 users. The Allworx servers support traditional analog telephony in a circuit-switched fashion between its analog FXO loop interfaces and FXS analog extension ports. It also acts as a VoIP gateway to bridge the digital data packet-based world and the analog world. As a result, the Allworx product supports ordinary analog telephones, analog telephone lines, and advanced VoIP telephones simultaneously —and *seamlessly*.

The Allworx VoIP technology platform is intended to integrate the following key features as smoothly as possible:

-
- *Auto Station Discovery*: Simply plug a new Allworx VoIP station into the network, and the Allworx server automatically discovers and configures the new station without manual intervention.
- *Simplify Move/Add/Change Administration*: With Allworx VoIP phones, station identity moves with the station, not with the physical cable drop used. This means that moving phones to new locations within the office is as simple as moving the phone to the new location and plugging it into another network drop.
- *Remote Phone Capability with Remote Plug-and-Play Functionality*: Install an Allworx VoIP phone at a remote site (e.g., home) and seamlessly operate it as if it was directly connected to the LAN at the office.
- *Site-to-Site Toll Bypass Calling*: Three or four digit dialing from one Allworx system to another Allworx system at a different site, without placing a call on the regular public phone network.
- *Integration with Internet Telephony Service Providers (ITSP)*: The Allworx server acts as a proxy server, placing and receiving low-cost Internet telephone calls through ITSP providers without requiring the use of regular phone lines. Allworx has performed interoperability testing with many of these ITSP' and a complete list can be found at www.allworx.com
- *Third-party SIP Gateway Product Integration*: In addition to the Allworx Port Expander (P/x 6/2) Allworx can expand via LAN-connected and SIP-based FXO, FXS, and/or E1/T1/PRI third-party gateway products.

4 SIP Protocol and VoIP

The Allworx VoIP platform is built around the industry standard VoIP protocol known as **S**ession **I**nitiation **P**rotocol (SIP). SIP is a packet-based protocol built on top of the standard IP stack using the User Datagram Protocol (UDP/IP). Although it is possible for VoIP telephony to use other standards (e.g., H.323 or MGCP), SIP was specifically designed for IP stacks, and was developed with Internet protocols in mind. While historically much of the older installed base of packet-based telephony used MGCP or H.323, it is generally accepted that SIP represents the future of VoIP – and nearly all new installations using industry-standard protocols are deployed using SIP.

When it comes to IP-based VoIP, SIP is not the whole story —it actually describes only one of the three functional elements required. In other words, when designing a VoIP protocol, three basic functions must be provided:

1. Call Control and Setup/Termination – A set of mechanisms to locate the intended dialed parties, determine their availability, and accept or deny their requests.
2. Session Negotiation – Once a new call is going to be accepted, this determines the format and network locations for transporting audio between the actual end points.
3. Media Transport – Once accepted and negotiated, this provides real-time audio transport between the end-points for the call's duration.

SIP itself actually only provides the first item described above. Other protocols are actually required to perform the other two primary functions. When people talk about SIP, two protocols are also generally implied:

- Session Description Protocol (SDP) to negotiate the session media types (G.711 or G.729) and the IP address and port number that each end-point should transmit toward.
- Real-Time Transport Protocol (RTP) to actually move coded audio data during the live call.

Therefore, when people talk about SIP VoIP telephony, several things are implied to be available and working properly for successful phone calls:

- Reliable IP routing data connectivity of UDP packets between associated phones and their gateways or proxy servers for basic network transactions (network settings, DHCP, DNS, etc.).
- SIP protocol and proxy configuration to locate intended parties and determine their availability (ringing or busy).
- SDP negotiation to determine the final coder type, IP addresses, and port numbers that should communicate actual audio data.
- RTP to transport coded audio over a network with an acceptable QoS level from end-to-end.

Acting as a VoIP gateway, the Allworx server contains all the necessary facilities to make the above happen in as simple a manner as possible. However, when one of these mechanisms is interfered with on the network, certain types of symptoms may result – such as dropped calls, choppy audio, one-way audio, or echo. Installers and site administrators must be certain that things are configured properly in their environments to ensure proper data connectivity and QoS between end points.

Looking forward, the remainder of this document will be dedicated to helping administrators and installers better understand the potential pitfalls, so that they may troubleshoot and resolve networking or configuration difficulties.

5 Echo in VoIP Networks

From time to time, echo can be a problem during telephone calls. Certainly everyone has experienced echo at one time or another, even for ordinary, non-VoIP calls. Echo is most commonly experienced in international calls or long-distance calls to rural areas, and of course in calls on cell phones. Unfortunately, the characteristics of VoIP telephony connections increase the opportunity for echo problems to occur. This is due to the introduction of additional latency (delay) of the voice as it travels from source to destination over the packetized data network, and is most acute when installed with analog lines.

To the human observer, echo of his or her own voice is only noticeable when it is heard back with some amount of delay. Echo without delay simply sounds like side-tone. Side-tone is the sound of your talking fed back directly between microphone and speaker without any delay. It is normally introduced purposely on phones so that the phone does not sound “dead” when you are talking. Therefore, when echo exists in the analog phone network, especially in local calls, it is completely covered up by the existing side-tone. However, when a VoIP system is attached to that very same analog phone line, the additional latency of the data network now carrying the voice to/from the IP phone now makes that echo noticeable to the user – since the echo now arrives back well after the speaker has finished each sentence.

Echo round-trip delays of only a few milliseconds (ms) are not really noticeable, but as the delay accumulates into the area of 10ms, the system will start sounding hollow and eventually will start sounding like the reverb of a large stadium echo. As echo latencies run into the area of 50ms and beyond, the speaker’s own speech will be followed by a very distinct echo of the same words back in his or her ear. At this point, unless the echo is very quiet, it starts to become annoying.

Generally speaking, the loudness and the latency impact how objectionable the echo sounds to the talker. As the echo gets quieter and/or less delayed, it becomes less objectionable. Alternatively, we can say that the more delayed the echo is, the quieter it needs to be in order to be acceptable. For echo to be acceptable in a VoIP system, it typically needs to be quieter than it was to start with in the analog-only part of the network. This is the role of the *echo canceller* in a VoIP system, which we will cover later.

5.1 Where does the echo come from?

Echo results in the phone network when two-wire phone lines carrying voice in both directions on the same wire pair are converted into four-wire circuits (where a separate wire pair carries the audio in each direction). The analog device that does these conversions is called a *hybrid* – and its job is to convert back and forth between the two-wire analog loop world and the four-wire Central Office (CO) switch world. If *hybrids* were perfectly matched to the particular phone and phone lines installed at every site, there would be no echo. However, in the real world, hybrids are not installed with perfect impedance matches, and therefore echo results when sounds “bounce off” the hybrids.

In a typical analog phone call, there will be at least two hybrids involved: one at the CO for the calling party (near-end echo) and one at the CO for the called party (far-end echo). Beyond the hybrid *electrical* echo, there can be other sources of echo. The most common is *acoustical* echo at the far end – when you call someone who has a low-quality phone or is using a desktop-type speakerphone. The near-end echo is determined by the local phone line loop on which you are making the call, while the far-end echo depends on the party being called. For this reason, near-

end echo is generally referred to as “line echo” and the other sources of echo are collectively referred to as “network echo”.

Delay times for “line echo” are not a concern for ordinary analog phone calls because the delay path is short enough that the echo sounds like side-tone. This means that the phone company can ignore the effects of line echo. However, the delays inherent in “network echo” are typically problematic, even for regular phone networks, especially with calls that cover long distances or have long loops from the telephone companies Central Office (CO). The phone companies have to do something about this – so they have devices built into long-distance phone networks called **Network Echo Cancellers** (NEC) to remove this echo. Now it is generally safe to assume that “network echo” is not a particular concern for long-distance calls, even with VoIP systems attached to the PSTN. Another noticeable difference is that on every analog connection, the user will hear what is known as “comfort noise”, a low background noise that confirms the connection is still active. With either digital or IP calls, no comfort noise is artificially introduced into the conversation so while the purity of the call is superior, any echo is not masked by the level of comfort noise.

5.2 The need for Line Echo Cancellers in VoIP

As we stated above, the PSTN is not concerned with line echo, since it will sound like side-tone. However, if we attach a LAN VoIP system to a PSTN gateway device (like Allworx), line echo becomes a specific concern in the system – because the hybrid echo coming back from the line is now delayed by tens of milliseconds in the IP network and will no longer be acceptable to the VoIP phone station user. VoIP gateway systems employ a device in their line interfaces called a **Line Echo Canceller** (LEC) that can cancel up to approximately 16 or 32ms of echo resulting from the hybrid installed on the CO’s local phone line.

5.3 If VoIP systems have an LEC and the PSTN has NECs, why do we still hear echo at times?

This is a complicated question with several different answers:

- An echo canceller is a very sophisticated device that automatically attempts to dynamically detect, adapt to, and remove all echo “on the fly,” while still providing true full-duplex speech performance. Neither LECs nor NECs are perfect devices – and depending on the design trade-offs of a given implementation, they will exhibit certain strengths and weaknesses in some operating environments.
- The phone company NECs never perfectly converge down to zero residual echo. When a VoIP system is introduced at one end of the connection, the increased delay may make the existing residual echo more perceptible. As described previously, this added delay gives the *perception* that the echo is worse, even though the magnitude of the echo signal is actually the same. In a given call, depending on the exact level of residual echo, this may or may not end up being objectionable to the VoIP system user.
- Regional “intra-LATA” calls can be very problematic relative to network echo. Because latency in the network is not significant, the phone company doesn’t usually bother to deploy the relatively expensive NECs for intra-LATA calls. The delay is relatively short, so the echo is not usually objectionable when using an ordinary analog phone at each end. However, add a VoIP system to one end and the network echo can be a real problem. This most often occurs when placing short-haul calls between competitive local or regional companies, and the called party has a particularly high level of far-end echo coming back. The inherent latency of the echo falls into the hole between the LECs ability to

combat the echo and the lack of an NEC in the phone network. To be clear, that echo was always there – it just took the VoIP system to actually hear it.

5.4 Shouldn't a VoIP gateway then have both an LEC and an NEC?

Deploying both LECs and NECs in VoIP gateways or PBXes has some advantages. In particular, it can help with regional calling area calls (intra-LATA calls), which are typically the most problematic for VoIP. However, intra-LATA calls are a small percentage of most users' calls, and having the NEC operating for other types of calls – the ones that already sounded good - presents some problems. The fundamental concern is having two different NECs operating on the same call: the NEC in the VoIP gateway and the one on the phone network. It can be difficult to get them to work reliably together, and in many cases it does more harm than good. You run into a situation where maybe five or 10 percent of calls are improved quite a bit, but the remaining 90 percent actually get a little worse. Which is a better trade-off?

In the end, most VoIP systems installed as end-user customer premise equipment, (including Allworx) employ only LECs, not NECs. The reasons are both technical and economic. NECs require significantly more processing power than LECs and are historically very expensive to implement on several channels at a time. Market forces have shown that the added costs are not outweighed by the functional benefits.

6 Conclusions

Other than new skills for both installers and administrators relative to data network Quality of Service (QoS) that are discussed in the next sections, echo is the biggest hurdle for VoIP systems to overcome, as they improve with each generation. Echo cancellers go a long way toward maximizing the user's perceived audio quality, but still represent one of the areas for the relatively new VoIP technologies to improve. Looking forward, the quality and capabilities of echo cancellers will continue to improve, but the only thing that will completely eliminate echo sources is when VoIP systems no longer need to interact with analog phone loops that date back to the designs of Alexander Graham Bell. Once calls between all end points are completely digital, the problem of occasional or persistent echo will be a thing of the past.

7 BANDWIDTH CALCULATIONS

This section provides details for a technical foundation of bandwidth calculations in Allworx-supported VoIP applications across the LAN and WAN. Also included are deployment recommendations that may help resellers and end-users with their application rollouts.

- In a VoIP telephone call, the caller's voice is converted to electrical signals which are then coded into data network packets or traffic. The coding/decoding (codec) scheme and the packet transmit interval collectively determine the amount of bandwidth consumed per call.
- G.711 calls send and receive a stream of Ethernet frames, each 214 bytes long, at a rate of 50 per second (20ms interval). The bandwidth required is 85.6kbps in each direction of the call. Calls that traverse Frame Relay, ATM/DSL, PPPoE, or VPN links will consume more bandwidth due to the additional encapsulation of the transport protocol(s).
- G.729A calls send and receive a stream of Ethernet frames, each 74 bytes long, at a rate of 50 per second (20ms interval). The bandwidth required is 29.6kbps in each direction of the call. Calls that traverse Frame Relay, ATM, or VPN links will consume more bandwidth due to the additional encapsulation of the transport protocol(s).
- Both G.711 and G.729A codecs may be used for VoIP calls through the Allworx. This typically involves a call from a VoIP phone through the Allworx server to another VoIP phone or soft-phone. Codec preference settings at each endpoint will determine which is used. For Allworx IP phones, this setting is configured on the handset's pages for each particular station.
- Only the G.711 codec is supported for calls in which the Allworx server is an endpoint. This typically involves a call from an Allworx-attached analog phone to another phone (VoIP or analog), a call from a VoIP phone to the Allworx voicemail system or auto attendant, or a call from a VoIP phone to an external user over analog public telephone network connections.

7.1 Codec Support Matrix

Calling/Called Endpoint	Codec Supported
Cisco 7905, 7912	G711
Cisco 7940, 7960	G711 & G729A *
Allworx 9102, 9202	G711 & G729A *
Allworx 9112, 9212	G711 & G729A *
Allworx 9224	G711 & G729A *
Analog set Analog telephone line	G711
3 rd party gateways	G711 & G729A **
Auto Attendant	G711
On-Hold Music	G711
Voicemail	G711

* Depends on preference settings and capabilities of the other endpoint.

**Subject to the capabilities of the third-party equipment.

7.2 Capacity Planning

Using the worst-case measurement of available bandwidth from above, you can calculate the maximum number of simultaneous calls supported over the Internet connection. This assumes no other Internet activities are being performed by local users (e.g., Web surfing, e-mail, file transfers, music downloads). Moderate to heavy use of the Internet connection for other applications will degrade the quality of calls and may substantially limit the number of calls supported over the link.

Available Bandwidth	Simultaneous G711 Calls	Simultaneous G729A Calls
128K	1	4
256K	2	8
384K	4	12
512K	5	
768K	8	
1M	11	

7.3 SIP PROTOCOL AND NAT FIREWALLS

The SIP protocol was designed and first implemented before security issues and the necessity for NAT/Firewalls existed. VoIP applications and the associated SIP and SDP protocols were not designed with **N**etwork **A**ddress **T**ranslation (NAT) in mind. In fact, SIP/SDP negotiations are typically broken when a NAT device exists between the negotiating end-points, so the resulting audio is not available in one or both directions after a call is set up. While a full discussion of this topic is beyond the scope of this document, the glossary at the end has a brief description of NAT. Relative to the effect of NAT on VoIP protocols, an understanding of the basic problem is useful, and is the topic of the remainder of this section.

NAT actually interferes with several different common protocols – SIP/SDP is only one pair of them. NAT breaks almost all protocols that need to embed IP addresses and/or port numbers in their own protocol messages. This is best explained through an example: let’s assume two phones are trying to talk to each other over the Internet. Each phone is behind its own NAT firewall at two different sites and the LAN network addresses of both sites are 192.168.1.0 with subnet masks of 255.255.255.0. When a call is setup, each phone is going to report through its SDP information an address of 192.168.1.x (its local IP address and port number) to the remote party. However, neither end is going to be able to send to the SDP-reported address and get the intended recipient. In fact, this will be a problem if the LANs had used the same network address or if they used any non-publicly routable network address. For this example, audio will not flow properly in either direction using normal SIP/SDP negotiations.

More typically, only one end is directly behind a NAT device, such as when contacting a remote VoIP gateway that connects to the PSTN. In these cases, audio typically works in one direction but not the other. There are also problems beyond the basic logical NAT routing issue, even if the IP addresses were right. The firewall function of the NAT introduces other packet filtering problems – since the whole point of the firewall is to prevent arbitrary packet data from entering the LAN network. To a simple firewall that’s not tracking all the SIP/SDP sessions going back and forth, the

audio data coming from the remote is simply blocked as “illegal data,” even if it did manage to route from end-to-end.

Since both NAT/Firewall and VoIP services are desirable, what do we do about SIP and NAT? There are several pieces required to get this to work correctly. The key element is a special device called an **A**pplication **L**evel **G**ateway (ALG). This is traditionally a separate type of special NAT firewall that is “SIP-aware.” It is specially configured to monitor the context of everything going back and forth between end-points – altering SIP/SDP/RTP packets and opening/closing holes through the firewall to allow the audio to negotiate and flow correctly.

Alternatively, you can use Allworx as your VoIP gateway with Allworx phones. They work together to solve all these problems for you automatically – even when using third-party firewalls. The details of this are explored in the next section.

8 Allworx Solves the SIP NAT Problem

Using VoIP protocols with NAT/Firewalls can be a big headache, unless you are using Allworx equipment. Allworx products are designed to work together – automatically discovering networking topology between end-points and adjusting all VoIP negotiations accordingly. Allworx products are able to do this even when third-party firewalls are involved in the path, but this requires the use of Allworx IP phones or servers at the associated end points. Non-Allworx end-points may not support all the necessary mechanisms to make this possible.

Generally speaking, the primary requirement is for each Allworx server to have its WAN port connected directly to the Internet at a publicly routable IP address. The Allworx server does **not** have to be the primary data NAT/Firewall for the LAN; it must have a publicly routable WAN connection in parallel with an existing firewall.

Note: First introduced with Allworx server software Release 5.2, many times it is now also possible to place the Allworx server behind a 3rd party NAT/Firewall. New advanced features were added in Release 5.2 to allow remote VoIP connectivity with a properly configured 3rd party firewall. See the admin manual for details.

Remote end-points (such as Allworx IP phones) on a LAN can typically be behind any single NAT/Firewall, whether it is an Allworx server acting as the firewall or any third-party NAT/Firewall product. Allworx specifically tests Cisco/Linksys™ and Sonicwall™ products as base verification reference points of “typical” firewalls. Since Allworx doesn’t control the implementation of third-party products, it can’t guarantee proper operation, but would expect proper operation with most firewalls, including the ones Allworx tested.

9 QoS on the LAN

This section relates to configurations where both the Allworx server and IP phones are located at the same physical site. Normally, all devices on the site will be on the same LAN subnet, so routers are not needed for internal LAN connectivity.

Strictly speaking, the textbook-recommended configuration for such a site is to configure a pair of VLANs using managed Ethernet switches. One VLAN would be for voice traffic and the other would be set at lower priority for data. Such a configuration basically guarantees that any amount of data traffic loading or problems cannot interfere with voice traffic. Allworx phones and their built-in switches have VLAN support to integrate with such a configuration, should it be desired. However, going to a fully managed VLAN configuration on a small business site is rarely done because of the administration complexity and cost. In general, following these tips will allow things to work well:

- Ensure the Local Area Network (LAN) is free of legacy hubs or repeaters and coaxial cable network segments. A completely modernized network with fully 10/100 switched Ethernet infrastructures is ideal.
- Minimize the number of Ethernet switches installed in the closet. Daisy-chaining together small switches to add more ports also adds latency and increases traffic flow bottlenecks. Installing one 48-port switch is much better than installing four 12-port switches.
- Group all VoIP devices onto the same Ethernet switch, if possible.
- The vast majority of Allworx installations do not require queuing capable switches or routers on the LAN. Both Allworx and VoIP phones employ sliding packet buffers that mask the modest packet loss and jitter (variable delay) associated with busy LAN networks. Managed routers and switches are only concerns in large enterprise networks.
- There is rarely a need to configure a VLAN setup on the network switches unless the customer's LAN is very large or users are extremely heavy data users. In these cases, the Allworx should be configured as a LAN host and a voice VLAN should be built on a separate switch to handle telephony traffic.

The simpler the site, the better the above tips will work. Things start to break down when network data traffic is regularly very heavy and the network is getting overloaded, or the site has onsite routers to direct traffic between more than one local subnet. If voice traffic is going to flow through those routers along with data, significant attention to QoS topics will be required to ensure proper operation 100 percent of the time. Although a full discussion of this case is beyond the scope of this paper, the next section talks about QoS issues over a WAN and addresses some of the issues involved.

10 QoS across a WAN

This section relates to configurations where the Allworx server and one or more IP phones are located at different physical sites. It also applies to cases where multiple Allworx sites are connected together in a site-to-site manner or when the Allworx server is configured to take advantage of an Internet Telephony Service Provider (ITSP) for calls to/from the Allworx server.

QoS topics across a WAN are of particular concern for both physical and historical reasons. The historical part of the problem is that IP protocols and the Internet in general were originally engineered to move only data – all treated pretty much equally on best effort basis. It didn't matter if the data was e-mail or coded voice traffic. As it sits today, there is no standardized way for the public Internet to support prioritized traffic between arbitrary end-points. Those protocols are still evolving, and the installed base of Internet infrastructure is not fully equipped to support the protocol standards, even where they do exist currently.

The basic physical problem here is what to do when bursts of data exceed the bandwidth of a limited size pipe. This is a complex topic that has several aspects. For example, there are priority and traffic shaping trade-offs that affect both the effective latency and available throughput of different traffic classification types. The traffic patterns and needs of various sites are different and have to be managed with site-specific knowledge of policies and priorities.

In many circumstances, ordinary Internet connections carry voice traffic pretty well most of the time. For most users, the potential reliability disadvantages are greatly outweighed by the cost advantages of a simple ad-hoc WAN setup. While guaranteed operation is only possible through a carefully engineered and managed QoS plan, adhering to the following guidelines will pave the way for a cost-effective solution using only ordinary Internet connections that may already be in place:

- Do not attempt to deploy VoIP service using a dial-up connection – these are too easily overloaded by even modest data traffic.
- For remote telephony applications to work through Allworx, the Allworx server may require to have its WAN interface directly connected to the public Internet. This is discussed in more detail in previous sections. In particular, calls to/from an ITSP service will typically not work if the Allworx server is behind a firewall.
- Prior to deploying VoIP between two sites, it is highly recommended that you first test your Internet connection to determine the speed of your link. Test the speed several times per day over the course of a week, and base your planning on the slowest rate measured.
- A VoIP call consumes symmetrical data on the network. Be sure your speed test results account for uplink and downlink performance. The lesser of the two values should be used for bandwidth planning.
- Determine the percentage of available bandwidth to be used for voice. Generally, it is better to not use more than about 50 percent of the available bandwidth for voice, leaving the remaining 50 percent for data applications.
- Compute the maximum number of calls your voice-allocated bandwidth will support and configure Allworx VoIP server settings to limit the maximum number of calls accordingly (so the desired limits are not exceeded). See Section 7.2 for a reference table.

- Use your local ISP's speed test if the remote application will traverse the same provider's network. Check your ISP's home page for a speed test link or visit BroadbandReports.com for a comprehensive list of 216 global sites to test your bandwidth <http://speedtest.broadbandreports.com>.
- Test the speed to the remote user's ISP if they use different providers. This will expose the performance of the peering connection between ISPs and provide a better perspective of the bandwidth available for the application.
- Use the free Brix Networks test utility for VoIP for Internet connection assessments. This will test your Internet connection's ability to handle VoIP calls. It also gauges the quality of the call in comparison to traditional and cell phone call qualities. Check them out at <http://www.testyourvoip.com/>.
- Allworx does support QoS tagging of voice traffic. However, it will make little difference in the caller's experience over the Internet with normal ISP-based services. Bandwidth availability should be the main concern today because ITSPs do not manage call quality to the customer using QoS features. The ITSPs that manage call quality only do so within their core networks; quality to the customer through ISPs is considered a best effort and will not be managed toward your network unless you subscribe to a dedicated private service with a specific Service Level Agreement (SLA) in place.

11 Key System Feature – Allworx BLF Protocol

The Allworx server and IP phones (local or remote) have several differentiating features that enable emulation of classic Key system-type capabilities – such as line appearance, busy lamp field monitoring, and direct station selection. This is all done on an industry standards-compatible SIP VoIP platform via the addition of some Allworx advanced mechanisms built on top of SIP.

Specifically, Allworx has added some SIP specification-compliant private headers to the Allworx implementations of SIP to activate the advanced features. These features include support for NAT-enabled remote phones and operations like automatic off-hook for direction station selection. However, the live system status monitoring required for things like busy lamp field indicators and line appearances goes beyond what SIP was designed for. As a result, Allworx developed a companion protocol for SIP to offer some of these advanced features. This is referred to as Allworx Busy Lamp Field (ABLF) Protocol.

The internal syntax of the ABLF is not important here, since Allworx takes care of the details. But it is helpful for the system administrator to have some understanding of how ABLF operates at the IP level. This will assist with troubleshooting specific problems where ABLF or line appearance lights do not seem to operate properly or consistently on a particular phone or set of phones.

11.1 ABLF Protocol Background

The ABLF protocol is an event-driven, peer-to-peer protocol that is implemented by both the Allworx server and Allworx IP phones. Because ABLF is a peer-to-peer protocol that uses sub-net broadcasts to reach all devices simultaneously, all ABLF devices associated with a particular Allworx server site must transmit and receive on the same UDP port number. By default, this UDP port number is 2088, but is configurable on a site-by-site basis using the “Servers -> VOIP Server” page of the Allworx administrative Web site to change the ABLF port option. When this setting is changed, all Allworx devices (IP phones and server) must be restarted to acquire the new configuration information and keep all devices monitoring and sending with the same correct port number.

11.2 ABLF Protocol Operation

Each time an ABLF peer has a change in status, it broadcasts this information to all peers on its subnet. If a peer is not located on the same LAN subnet as the Allworx server, the ABLF peer also sends a directed packet with the same content to the Allworx server so the server can forward that packet to all other subnets with phones attached. If the phone is an Allworx remote phone, this directed packet goes to Allworx’s WAN address, otherwise this is typically directed at Allworx’s LAN port, since the device is on the private side of Allworx’s firewall. The ABLF device automatically determines the interface to use when it is configured during startup.

The Allworx server keeps track of a list of all local subnets that contain ABLF devices and also contains a list of each remote device participating in the ABLF protocol. When the server receives a packet from any device, it automatically forwards the packet to a single device on every subnet (other than the originating one) and requests that one device to broadcast the packet on its own local subnet. This mechanism lets the subnets learn the topology of all ABLF devices and manages the traffic so that every device gets each ABLF update notification in as efficiently a manner as possible.

11.3 Troubleshooting ABLF

Generally speaking, the ABLF protocol and Allworx devices take care of themselves, even when NAT/Firewalls exist at the remote sites sitting in front of the remote phones. This is because the Allworx server and phones work together to safely traverse NAT devices within the same restrictions explored previously when talking about SIP. It is not even normally necessary to open a specific NAT firewall device port at the remote site to support ABLF – the phone will keep an automatic port open with the server. In the event that one or more devices or subnets are not correctly receiving ABLF packet updates, something is probably administratively blocking traffic on the configured ABLF UDP port (2088 by default). Make sure that nothing is specifically blocking port 2088 to the Allworx WAN port and that no intermediary NAT/Firewalls limit port 2088 traffic from local LAN to WAN. If port 2088 is limited or blocked, you can change the Allworx server's receiving UDP port as specified previously.

12 Remote Office Phones

The Allworx server and Allworx VoIP phones are designed to work together as seamlessly as possible. In particular, Allworx has tried to make it as straightforward as possible to install and maintain remote office phones – nearly as easy as local main office phones. In this section, “remote phone” refers to a standalone phone that operates in conjunction with an Allworx server located at distant site, without having a *local* Allworx server installed. The most typical example would be a phone at an employee’s home working off the main office’s Allworx.

In most circumstances when using server release of 5.2 or newer a single properly configured 3rd party firewall may reside between the Allworx server and the Internet if the public address assigned by the 3rd party device is properly configured into Allworx. However, in some circumstances, it will be important to have the Allworx server installed with its WAN port directly connected to a publicly routable IP address on the Internet. The exact requirement here is a property of the 3rd party firewall capabilities and whether or not any used ITSP’s will support the desired configuration.

The remote phone itself must have Internet access to the Allworx server, but the remote phone is typically allowed to be installed behind a single NAT/Firewall. This is usually the NAT/Firewall protecting the LAN on which the remote phone resides.

Configuration of the remote phone is mostly automatic and works similarly to installing phones on the main LAN of the Allworx server. The remote phone requires two specific settings to be set by the administrator:

Boot Server IP	The public IP address of the Allworx server’s WAN port. This IP address tells the phone where to find the main office server on the Internet.
Remote Plug-and-Play Key	This numeric code number is the authentication key that the phone uses to authorize itself for the remote Allworx server. If the correct key is not entered, the Allworx server will not allow the remote phone to place calls through it. This is a security measure to help prevent unauthorized users from using services on the Allworx server. The correct setting is located in the configuration setting on the Servers->VoIP page of the server’s administration screens. The value is common to all remote phones associated with a particular Allworx Server.

Beyond this, generally speaking, the remote phone operates the same as any other phone on the main site’s LAN. The station can make and receive calls and be configured just like any other station. There are only a few limitations in capabilities:

- Intercom works fine, but paging functions do not extend beyond the local LAN. Neither zoned nor overhead pages will typically play out at the remote site.
- If more than one remote phone is behind the same third-party NAT/Firewall, the remote phones **may** not be able to call each other successfully without some specific 3rd party firewall configuration.

- Remote phones off a single Allworx server at different sites should be able to call each other, but if both connected phones are behind NAT/Firewalls, getting audio between them can be problematic. It is sort of a “chicken-and-egg” type problem: neither phone can determine its public IP address or RTP port numbers until after the opposite end receives the first audio packet. Since both ends are in the same situation, neither end ever receives the first packet. To resolve this, at least one of the two phones must have a static mapping through the firewalls for its RTP ports. See the troubleshooting section for more details.
- ITSP services configured on the Allworx server may not be accessible from remote phones that are behind NAT devices. This is a limitation of the service providers and not under control of the Allworx devices.

Caveats and details about remote phones operating over the Internet, especially when using third-party firewalls:

- Normal ISP Internet access for regular residential and business customers is a best effort service. Specific QoS metrics are not guaranteed and poor quality audio can result at times, depending on traffic flow between providers and through Internet peering points.
- Most low-cost NAT/Firewall routers do not prioritize traffic, and even if they do, the Internet service that they are being used with typically does not. Therefore, normal user data traffic activity can affect audio quality. For example, downloading e-mail while talking on the phone may cause audio interruptions.
- Depending on the service provider and the quality and bandwidth of your ISP service, the above issues may be rare or they may be occurring regularly. See subsequent sections for more details.
- If none of the above are acceptable from time to time, then leased line and/or virtual private circuit-type service with a specific Service Level Agreement (SLA) is required – along with proper router equipment – to guarantee that the desired QoS metrics are always met.

13 Zoned Paging

Allworx system paging features use a special form of IP traffic routing called “*multicasting*” to direct zoned pages from the Allworx server to all Allworx phones configured for a particular zone. Multicasting is used because many phones need to receive the page simultaneously and this is exactly what Ethernet and IP multicasting were designed for.

Phones subscribe to a multicast group to become a member of a particular paging zone, since each zone has its own multicast address. Generally, this is all automatic and transparent to the end user, the administrator, and even Ethernet switches. However, when a particular site is having trouble with pages reaching one or more phones, understanding the configuration options and how Allworx transmits these pages should be helpful in diagnosing site configuration problems.

Allworx zoned paging is configured on the “Servers > VOIP Server” page of the Allworx administrative Web pages. In particular, there are three parameters used to determine where Allworx transmits zoned pages:

Configuration Item	Description
Paging Base Multicast IP Address	This is the multicast base IP address (zone 0 - the overhead zone) of the system. Follow-on zones are numbered sequentially from this base by adding the zone number to the base IP address. For example, zone 5 is located at base + 5 multicast address.
Paging Port Number	This setting is the UDP port number to which the server transmits at the multicast IP address for the current zone. All zones use the same port number, each with their own multicast address as described in the above field. The actual UDP payload is RTP packets of 20ms G.711 frames.
Paging Max Hop Count	This value controls the time-to-live count value in the IP header of all paging UDP/RTP frames. This value is usually set to one (1) for a single LAN subnet, but if you have multiple LAN subnets with phones on them, this value may need to be increased.

Generally speaking, pages are only implemented on a single site and a single LAN subnet. However, by manipulation of the above parameters and through configuration of intermediate routers, it is possible for Allworx pages to span multiple subnets and even to tunnel across a VPN between sites. What needs to be configured on the site is a way for Allworx server-sourced multicast packets to be routed between subnets using the normal router-specific configuration mechanisms. A full discussion of routing multicast IP packets and how this is configured into different brand routers is beyond the scope of this document.

14 Common Problems and Tips

This section explores some commonly observed problems and possible causes not specifically addressed by the previous sections. This section also provides some general troubleshooting tips.

14.1 One-way audio

Having audio flow in one direction only after placing or receiving calls, or even having **no** audio flowing at all is a common problem with SIP and RTP, especially when NAT/Firewalls are on the path. As described in previous sections, Allworx products implement several mechanisms to automatically deal with the majority of configuration complexities required to get SIP and RTP to work with NAT firewalls. Still, it is possible for specific routing asymmetries or packet filtering rules to interfere with one or more protocol mechanisms.

The most common configuration problem in this regard is in complex setups, where the Allworx server is connected to the WAN and another firewall on the LAN is used for normal data traffic. The user-level symptom is that a LAN phone will not get good audio either inbound or outbound to a remote phone located at another site. While one might conclude this is a remote site configuration issue (since the LAN phone works normally when talking to other LAN phones), this is not the case. Generally, the LAN gateway is set in the phones so that Internet traffic traverses through the non-Allworx server firewall that is not SIP/NAT aware. In these cases, it is important to configure the LAN IP phones so that they use the Allworx server LAN IP address as their gateway to the Internet. This allows the Allworx to correctly orchestrate firewall filtering to pass the remote phone audio traffic to the LAN properly.

Note: As of Allworx Server Release 5.2 and later, Allworx IP phones will automatically detect this configuration concern and adapt automatically without having to manually set a unique gateway setting in the phone. However, 3rd party phones may not have this capability, so one should still be aware of this potential concern.

14.2 Intermittent connectivity or devices dropping off the network

In cases where connectivity between devices on the network or between IP phones and the Allworx server is spotty or random, some sort of configuration or network topology problem usually exists. The primary things to look for are duplicate IP addresses on the network, multiple DHCP servers enabled on the LAN, or more fundamental QoS issues (including bad cables or improperly configured VLANs). It is normally best to start the investigation at the point of recent changes to network topology or on the newest configuration settings.

Connectivity problems, especially when intermittent, can be very subtle and sometimes difficult to track down. A network packet sniffer or protocol analyzer is a great companion to analyze packet flow and look for problems. If all else fails, generally the best thing to do is isolate as many things as possible from the network and start adding things back one-by-one (over time) to discover which added device is interacting with the already existing devices on the LAN. Just because a specific device starts the trouble doesn't necessarily mean that it is at fault, only that it is a necessary accomplice. A careful review of packet data on the analyzer and/or configuration settings is always justified.

14.3 Common MAC address on LAN and WAN ports of a router or firewall

Most dual Ethernet interface firewall products (including Allworx) use the same MAC address for both or all interfaces of the device. This is perfectly legal, however there is one specific instance where this can cause traffic routing difficulties. When two dual interface devices are hooked together in parallel, both with a common routable path between networks, difficulties can arise. The most common case of this is when Allworx and some other router/firewall are hooked up with both the LAN and WAN interfaces between the same pair of networks (logically, the two WAN interfaces and the two LAN interfaces are connected together). This configuration causes a problem because some routing decision optimizations are done between layers two and three of the network protocol stacks. Because the same MAC address is reachable from two different interfaces and both devices have visibility to both paths to that address, neither device can determine the correct routable interface to use to reach the intended destination.

To prevent this from becoming a problem, you must configure the network so that the devices do not have visibility to both interfaces of the other. Generally, the easiest solution is to avoid the above configuration – or if it must be used, configure the WAN sides of these devices to be on different VLANs so that they cannot see each other's traffic. Assuming a hypothetical 3-port switch, put one device on port 1 (VlanId=1), the other device on port 2 (VlanId=2), and the Internet router on port 3 (VlanId=1&2).

14.4 DTMF digits not passed during live calls

Negotiation of how DTMF number key digits pass during live calls is the least standardized and most problematic aspect of doing VoIP telephony with SIP. With Allworx equipment, this will not be a concern, but it does often create a problem when operating with ITSPs. The most typical symptom is that PSTN phone calls coming in from an ITSP will not be able to operate the Allworx auto-attendant or voicemail applications. The reverse may also be a problem – where Allworx system phones will not be able to operate phone applications out on the PSTN. These problems happen because during call setup negotiations, the service provider may not have indicated correctly how it plans to send and receive DTMF digits.

Allworx products expect all SIP parties to negotiate DTMF out-of-band so that DTMF can work correctly with all coder types (including G.729). If you are having these problems, the default negotiation settings can be controlled in the advanced settings tab of the Allworx server SIP Proxy and Gateway configuration pages of the administrative Web site. You can control the default RTP type and whether padding of RTP payloads should be applied.

14.5 Music On-Hold sound quality is poor

Some voice coders are specifically designed to compress human voice signals – and **not** arbitrary audio content. This is case for G.729 and its variants that take full advantage of human vocal tract properties to get such levels of high-quality compression. However, when transmitting audio that is not speech (such as music) through these compressed streams, the resulting audio is distorted. Configure the system to use only the G.711 coder if this is a concern.

14.6 Calling Auto Attendant

When a user reports problems with calls to or from a particular station, or when calling a particular party, it is usually best to let the Allworx auto-attendant assist with troubleshooting. If a station has correct connectivity to the Allworx server, dialing extension 400 should always return your own familiar auto-attendant outgoing greeting in short order. If there are delays in connecting to the auto-attendant or a fast busy congestion signal is returned, then there are

connectivity or connection problems that must first be addressed. Additionally, once the auto-attendant is reached successfully, dialing “#7” will have the auto-attendant report back audibly which station it thinks you are calling from – so your station identity can be confirmed. When you hang up (per the auto-attendant instructions), the system will automatically place a call back to your station to confirm return connectivity to your location. If the system is properly configured and good connectivity exists, this auto-attendant mechanism should always work. It can be applied individually to each station to isolate the devices that have configuration or connectivity concerns.

14.7 Mapping ports through remote phone firewalls

When an Allworx server has remote phones at multiple sites and each site has a NAT/Firewall device between the phone and the Internet, the remote phone’s RTP ports must be statically mapped through the firewall if you want to call each other. It is not necessary to do this if remote phones only need to call the main office (and not each other). The steps to perform this operation vary by the type of firewall that the remote phone uses, but the basic goal is to map a range of UDP ports on the firewall’s IP address to the LAN address ports of the IP phone. The Allworx server’s admin page for each handset allows you to control the range of ports the phone will use for RTP sessions under UDP. On the firewall product’s configuration page, you need to map those same ports (one-to-one) through the firewall – thereby allowing access from the WAN to LAN through those ports. Some products call this “DMZing” those ports or “port mapping.” The phone’s LAN IP address must also be entered into this firewall configuration page, so is it wise to make sure this address will be fixed.

Glossary

AEC – Acoustic Echo Canceller. An echo canceller that cancels the effects of audible room echo, such as in a desktop speakerphone. See “Echo Canceller.”

Echo Canceller – An algorithmic device, usually implemented in software in a digital signal processor chip to remove the echo of the speaker’s voice from a full-duplex telephone connection. Echo cancellers are more sophisticated and differ from echo suppressors in that they eliminate echo without having to impose half-duplex communications on the link (like echo suppressors do).

Full Duplex – Allowing communication (including the movement of data or voice) in both directions of the circuit at the same time – simultaneous two-way communication. A half-duplex connection allows communication in both directions, but in only one direction at a time.

Jitter – The amount of variability in latency as a function of time. In VoIP systems, jitter describes the irregularity of packet arrivals over the course of time. As network jitter increases, deeper receive buffers are required to smooth the effects (and not lose packets). Increased jitter typically results in increased overall end-to-end latency.

Latency – A specific type of delay in time of transmission or response to events. More formally, is the amount of time elapsed between two fixed reference points in a system that transmits information or takes action based on certain events. Latency is sometimes thought of as the time delay between cause and effect. In VoIP, there are several sources of latency, including network latency, coder latency, packetization latency, buffering latency, and more.

LEC – Line Echo Canceller. An echo canceller that cancels the effects of electrical circuit echo in phone lines. See “Echo Canceller.”

NAT – Network Address Translation. Generically, NAT refers to devices that translate the IP addresses of packets as they transit from one subnet to another. NAT is normally associated with firewall devices and is used to hide private, non-routable IP addresses of a LAN from the public Internet. Using NAT has security advantages and also works to minimize the number of public IP addresses required at a single site. Minimizing the use of public IP addresses is important because the pool of available addresses is a scarce resource.

Side-Tone – Feedback of speaker’s voice from the microphone to the ear piece to make speech sound as natural as possible. Without side-tone, phone handsets sound like they are broken – even though speech is still being transmitted to the remote end.

Author: Jeffrey Szczepanski, Chief Technical Officer, Allworx, a wholly owned subsidiary of PAETEC Communications Inc.

© 2009 Allworx is a wholly owned subsidiary of PAETEC Holding. All rights reserved.. All other names may be trademarks or registered trademarks of their respective owners.