VoIP-based Air Traffic Controller Training

Supreeth Subramanya, Xiaotao Wu, Henning Schulzrinne

Department of Computer Science, Columbia University, New York, NY, USA

{sbs2139, xiaotaow, hgs} @ cs.columbia.edu

ABSTRACT

Extending VoIP beyond the Internet telephony, we propose a case study of applying the technology outside of its intended domain, to solve a real-world problem. This work is an attempt to understand an analog hardwired communication system of the U.S. Federal Aviation Administration (FAA), and effectively translate it into a generic, standards-based VoIP system that runs on their existing data network. We develop insights into the air traffic training and weigh on the design choices for building a soft real-time data communication system. We also share our real-world deployment and maintenance experiences, as the FAA Academy has been successfully using this VoIP system in 5 training rooms since 2006 to train the future air traffic controllers of the U.S. and the world.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design – *Network communications;* C.3 [**Special-purpose and Application-based Systems**]: Real-time systems; J.2 [**Computer Applications**]: Physical Sciences and Engineering – *Aerospace;*

General Terms

Design, Experimentation.

Keywords

VoIP Application, IP Multicast, Air Traffic Controller Training.

1. INTRODUCTION

Traditionally, voice communication has required dedicated infrastructures – both in the backbone and at the end-points. Then in mid-1990's, the ubiquitous connectivity of the Internet with its ever increasing bandwidth fuelled the research that made communication over the data networks possible (most notably on the Internet, referred to as Voice over Internet Protocol or Internet Protocol communication). Treating voice as a "real-time" data brought numerous advantages – reuse of the existing data network infrastructure that reduced the deployment and usage costs, digitization of voice that led to superior voice processing and compression techniques, and decoupling of voice service from the underlying infrastructure that fostered innovation in communications. Despite the widespread adaptation of IP communications in the corporate and commercial environments, the Internet and long-distance telephony over the last decade, some defense and federal agencies still continue to use the non-digital, custom-built traditional communication systems.

The FAA is an agency of the U.S. Department of Transportation, with the authority to regulate and oversee all aspects of civil aviation in the U.S. Its education and training division, the FAA Academy, provides technical training for the aviation community including future Air Traffic Controllers (ATC) and other aviation personnel in a variety of simulated and real environments. To conduct these trainings, the FAA Academy has traditionally employed three disjoint networks in every classroom and laboratory – a voice network, a graphic simulation network and a data network. The graphic simulation and voice networks are custom-built analog infrastructures to support flight simulations and inter-position communications respectively. Data network is a TCP/IP-based Gigabit Ethernet infrastructure with a firewalled connectivity to the public Internet.

The reasons for keeping these networks separate and not adapting seemingly superior technologies like VoIP are many. First, the system requirements are quite complex to readily reuse the components from the existing commercial VoIP software base, and the market is too small to attract vendor interest. Second, the uncertainty in migrating to a newer technology with its associated learning curve outweighs the trouble in putting up with the limitations of the existing system. Third, pressing short-term goals (e.g., meeting the demand for trained ATCs [1,2]) and limited financial resources delay system changes.

Columbia University CS Technical Report, September 26, 2008.

Figure 1. Low and medium fidelity training room[1]

Figure 2. High-fidelity training room[1]

Equipped with a legacy analog hardwired communication system, the FAA Academy faced a tough challenge in keeping up with advancements in the training programs – creating new training scenarios, accommodating additional students or just moving a student from one position to another required physical rewiring and reinstallation of the system. The voice system was getting prohibitively expensive, with its hardware needing custom manufacture. The FAA Academy then decided to collaborate with an academic research lab to seek a better alternative.

This paper discusses a novel case-study of a successful design and deployment of a VoIP system for the FAA. We gather and analyze the specialized communication requirements of the air traffic training domain and translate these into an IP-based design using standard Internet protocols. We describe our experiences and the lessons learnt in deploying and maintaining the VoIP communication system in five training rooms at the FAA Academy for the Initial Terminal Training, the Terminal Radar Training, and the International Training programs since early 2006.

Our work also illustrates the challenges of working in a conservative environment, where extremely high system uptimes are expected and the fact that with due diligence academia can very well create non-toy systems. We are not aware of any other real-world attempt to use IP communication in air traffic training systems.

The remainder of the paper is organized as follows. Section 2 introduces air traffic controller training, including the operational aspects, communication scenarios and the hardware interfacing. We analyze the design choices, describe the system architecture and outline the development methodology in Section 3. The performance analysis and system evaluations are presented in Section 4. Section 5 outlines the operational challenges in our real-world deployment and its influence on the system design. Section 6 provides a discussion of the related work and Section 7 concludes the paper.

2. BACKGROUND

2.1 Air Traffic Controller Training

Air traffic controllers (ATCs) are the personnel who operate the air traffic control systems to promote a safe, orderly and expeditious flow of air traffic. While at the FAA Academy, the future ATCs learn the techniques of managing air traffic – by developing mental pictures of air-space and air-time, by learning to communicate and coordinate with the pilots, the neighboring ATCs and facilities, and by learning to use a multitude of air traffic control displays and devices. The training is conducted in a variety of environments at varying levels of difficulty starting with the low fidelity (instructional games focusing on individual training), and moving on to the medium fidelity (real-time interactive training) and the high fidelity (complex interactions involving real hardware). Figures 1 and 2 illustrate facsimiles of these training environments.

The training rooms at the FAA Academy are designed to closely match the real-world experiences of an ATC, except that all ATCs are seated in the same room and that the pilots are not actually flying aircraft. In the high-fidelity training, every ATC is paired with a *position instructor*, who provides a continuous one-on-one mentoring while the course is taught by the *lead instructor*. In a typical session,

[1] The pictures are only indicative and do not represent the actual training rooms of the FAA Academy. Fig 1 and Fig 2 are from http://www.cba.uri.edu/classrooms/pictures/computerlab.jpg and http://www.lockheedmartin.com/data/assets/10307.jpg respectively.

after explaining the conceptual aspects, the lead instructor loads the relevant flight scenario onto graphical simulation and voice communication systems, instructs the pilots to simulate their flight movements, and then lets the ATCs and the pilots to interact with each other much like in real life.

2.2 Air Traffic Communications

This section briefly describes the different modes of communication modeled after the real-world interactions between the air traffic entities – the ATCs and the pilots. Also a few additional forms of communication, those involving classroom instructors, have been created specifically for the training environments.

2.2.1 Radio lines

Radio lines represent air-to-ground transmission, where all the transmitters and receivers tune into a selected frequency channel before communicating. There is no explicit signaling required for the communication, but simultaneous transmissions get garbled. The hardware and user interfaces provide the ability to use these channels in receive-only or send-receive modes. These are typically used for broadcasting messages to all the air traffic entities within a given sector. Sector refers to the air space managed by an ATC and all the sectors use the same set of frequencies, without interfering with the communications of the neighboring sectors.

2.2.2 Ring lines

These are bi-directional communication channels, similar to traditional telephony. When the caller initiates a call, both the caller and the callee hear ringing. The call gets established only when the callee explicitly accepts the call invitation by pressing a button, analogous to picking up the handset. Without this explicit acknowledgement from the callee, no channel is established. These are typically used for non-emergency one-to-one communication between a pilot and his controller.

2.2.3 Override lines

Override lines are bi-directional communication channels that do not require an explicit acknowledgement from the callee for the channel to get established. In other words, a call session gets created when the caller places a call and can only be terminated by the caller. In air traffic training, these are used for both bi-directional communications and passive overhearing of the neighboring controllers' communication, with the sidetone resulting from the bi-directional session establishment letting the callee know that he is being overheard.

2.2.4 Shout lines

Shout lines are the channels with an ability to support both uni-directional and bi-directional communications. When the caller initiates a call, a uni-directional media session gets established and the callee starts hearing the caller. If and only when the callee accepts, the media session becomes bi-directional. A uni-directional call can only be terminated by the caller, whereas a bi-directional call can be terminated by either party. In the air traffic domain, these are typically used for inter-facility communications.

2.2.5 Intercom lines

These are training-only channels, very similar to the ring lines in operation, typically used for inter-position communications. These are particularly useful in advanced trainings, where students working on the same training problem have to sit in different rooms or labs, due to the space constraints of large equipments.

2.2.6 Instructor monitoring/broadcasting

Instructor monitoring and broadcasting lines are training-only channels used by the classroom instructors (both position and lead instructors) for training-related communications. Monitoring channels are bi-directional in nature, not involving any visual cues at the student position, such that the instructor can passively listen to all the communications of a student and may choose to give occasional feedback. The broadcast channels are uni-directional in nature, typically used for classroom-wide announcements.

2.2.7 Classroom recording

For archival and logistical purposes, the FAA Academy records all the training sessions. Any recorded session can be played back in VCR-fashion, often in conjunction with the graphical simulation.

2.2.8 Notification of transmissions

Very often, the communicating entities, especially those who are not already tuned into a channel, need to know the current activity on that channel before starting to communicate. Such mechanisms that continuously provide a real-time feedback of the channel status are essential when channels have to be used exclusively (e.g., override lines) or to minimize crosstalk on highly collaborative channels (e.g., radio lines).

2.3 Air Traffic Devices and Interfaces

Air traffic training employs a multitude of devices but the communication system interfaces with only three types of devices – display devices, audio devices and channel selection devices.

Amongst the display devices, the students use touch screen monitors to interact with the communication system - to place or receive a call, to see the transmission status of the channels, and also to configure other hardware devices like volume of the speaker, sidetone of the headset, brightness of the display. These displays are typically 12" to 14" in dimension and similar to the ones commercially sold by ELO [3] and GVision [4].

Two audio devices are used - loud speakers and Push-To-Talk (PTT) headsets [5]. The PTT devices, usually manufactured by Plantronics, have a built-in sound card and support USB and serial interfaces. The students can choose either of these devices for receiving audio.

The students can listen simultaneously to many communication lines, but they can transmit on only one type of line at a time. The PTT and foot-pedal [6] are the channel-selection devices used by the students to select the transmission channels. Both of these provide a binary selection - a pressed state to transmit on radio lines and a released state for non radio lines.

3. DESIGN AND IMPLEMENTATION

The design goal is to create a voice communication system that makes use of the existing data network infrastructure, and supports all the functionalities present in the proprietary hardwired analog voice system. Additionally, such a real-time IP communication system should offer configurability (i.e., support for creating new scenarios by decoupling student roles from classroom positions) and programmability (i.e., support for adding new communication channels with different rules without any hardware-level changes), both of which were lacking in the hardwired system. Last, the new system should be at least as robust and reliable as the existing system.

3.1 System Architecture

Real-time communications comprise of two central aspects – signaling and media transfer. Signaling is the process by which the caller locates the callee, conveys her request to the callee specifying the communication type and parameters, and gets a positive or negative acknowledgement from the callee with her accepted set of parameters. After successful signaling, the communicating entities know exactly where and how to send and receive the media information.

3.1.1 Signaling

The signaling requirements for the air traffic communications discussed in the Section 2.2 are more demanding than traditional telephony, as every communication channel should be able to specify its own signaling logic. To this end, we decided to use Session Initiation Protocol (SIP) [7], an application-layer control protocol for creation and management of multimedia sessions and conferences that supports a flexible programming interface to signaling. An entity in SIP paradigm that represents a communicating endpoint is referred to as the SIP User Agent (UA), and accordingly our setup has two types, the student UA and the instructor UA. The UAs communicate with one another with the help of a SIP proxy server which is responsible for the user location, authentication and authorization, and request/response routing.

The signaling logic in SIP can exist in the SIP server or in the SIP UA. In our design, we have split the signaling logic between the UA and the server. The signaling component at the UA handles its behavior as a caller and a callee, for all the types of lines – including user's interaction with the I/O devices, audio-visual cues in the GUI, handling of incoming and outgoing call requests and responses. The signaling logic at the SIP server determines if the caller is permitted to make that call, and if so who should receive this call. We have adapted SIP-CGI [8] to input the service logic at the SIP server.

3.1.2 Media Transfer

An intuitive design for the media flow is a hybrid architecture where multicast is used for the radio lines and instructor announcements, and unicast for point-to-point communications. But when we consider the requirement that every possible communication should be able to be monitored, interjected and recorded, we realize that all communications could be effectively designed as multicast sessions or conferences. There are several ways of handling conferencing and audio mixing in VoIP systems and Table 1 compares the candidate technologies.

Table 1. Technologies for conferencing support

	Control	Scalability	Security	Implementation	Firewall Traversal
Centralized conference	Best	Fair	Best	Fair	Good
Cascaded conference	Good	Best	Good	Hard	Good
Full-meshed conference	Fair	Bad	Good	Easy	Fair
App. layer multicast	Bad	Fair	Fair	Hard	Fair
IP multicast	Worst	Fair	Worst	Easiest	Hard*

Hard*: Works with only firewalls that support multicast

In general, it is easier to achieve security handling and conference control, such as floor control and access control, at a centralized conference server because it has a single control point. In terms of scalability, since the cascaded conference distributes audio mixing computation to multiple conference servers while not introducing more network traffic, it offers the most scalable solution. In terms of implementation, the full-meshed conference, application layer multicast and IP multicast do not require a server implementation, but they all need the support of endpoint mixing, which is simpler than implementing a conferencing server.

To choose a solution, we first analyze the requirements and the deployment context of our system. First, there is no specific conference control functions required in our system, as the students themselves use Push-to-talk (PTT) devices to switch between listening and talking modes. Second, our system will be used in an intranet that is behind a firewall and will only communicate with entities inside the intranet. Therefore, we do not need firewall traversal capabilities. Third, since all the computers are located in the training rooms and all the users are the academy students, both the users and the communication endpoints are trustable and consequently, no specific security mechanisms are required of the VoIP system.

Last, we address scalability. A training room consists of no more than 27 communicating entities (26 students and 1 instructor in the low/medium fidelity training; 18 students and 7 instructors in the high-fidelity training). Given that in each sector a maximum of six radio lines and one point-to-point line can be simultaneously active, there can be no more than 91 conferences, and each conference will have no more than five participants (even when the communication is being both monitored and recorded). This is just a theoretical maximum, given that most training does not require all the students to be active on all the seven possible channels simultaneously. Even if later on, the number of students was to increase, it will only increase the number of conferences, but not the number of participants in a given conference. Since the least scalable of all solutions, namely full-meshed conference and the IP multicast, can support fewer than 10 participants in a conference, scalability is not a differentiating factor for our design choice.

Based on this analysis, we chose to use IP multicast, since it was the easiest to build among all the candidate choices. We have designed the media engine to employ Real-time Transport Protocol (RTP) [9] for packetizing and transmitting the digitized voice, and Real-time Transport Control Protocol (RTCP) [9] for getting QoS updates. Putting together all the required components, Figure 3 illustrates the system architecture. The dotted box in the Figure 3 indicates a logical grouping of an ATC student, his Position Instructor and two Pilots of that sector.

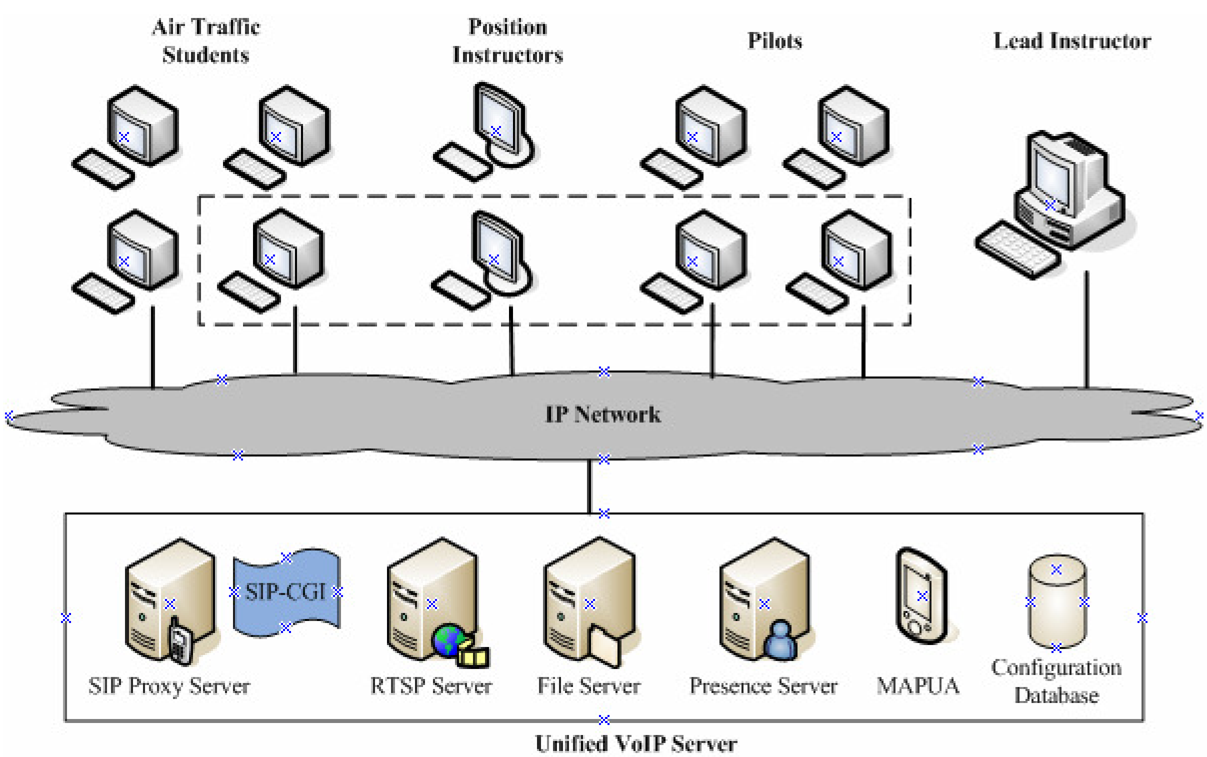

Figure 3. Communication system architecture of a typical training room.

3.2 Detailed Design

3.2.1 Radio Communication

Radio lines are essentially in-sector broadcast communication channels. Each radio-line is assigned a permanent SIP URI (like *sip:radio_120@faa.gov* for the 120 MHz channel) that can be used uniformly by callers in any sector. But internally, every radio-line in every sector has to be associated with a unique multicast address, to avoid interference across the sectors. This translation from the generic SIP address of a radio-line to its unique multicast address is performed dynamically by the Multicast Address Provider UA (MAPUA).

The call flow diagram for radio-line communication is shown in Figure 4. When a student UA calls a radio-line, the SIP proxy server executes the call logic associated with the radio lines, which redirects the request to the MAPUA. Then the MAPUA queries the configuration database to find out the exercise and the sector to which the calling student UA belongs. If the requested frequency for that particular sector and exercise has not already been assigned a multicast address, the MAPUA dynamically assigns a new address. Finally, the MAPUA accepts the call by returning the appropriate multicast address to the student UA, which in turn joins that multicast session.

3.2.2 Land-line / Inter-position communications

All the non radio lines are typically referred to as landlines in the air traffic parlance. Despite having seemingly different call behaviors, a central characteristic of ring lines, shout lines, override lines and intercom lines is that all are one-to-one communications. Thus, they can be conveniently grouped under a single umbrella, with the differences in their call behavior being handled by the call logic scripts. As noted earlier, even this apparently one-to-one communications are established as multicast sessions, to accommodate dynamic addition/removal of the instructors who may want to monitor the communicating students, or the recording agent, which may want to capture this communication. Similar to the radio lines, each land line is assigned a permanent SIP URI (like *sip:landline_mlcAFSS@faa.gov* for a ring-line called McAlister Flight Service) that can be used uniformly by callers in any sector or exercise. But internally, every land line in every sector has to be associated with a unique multicast address.

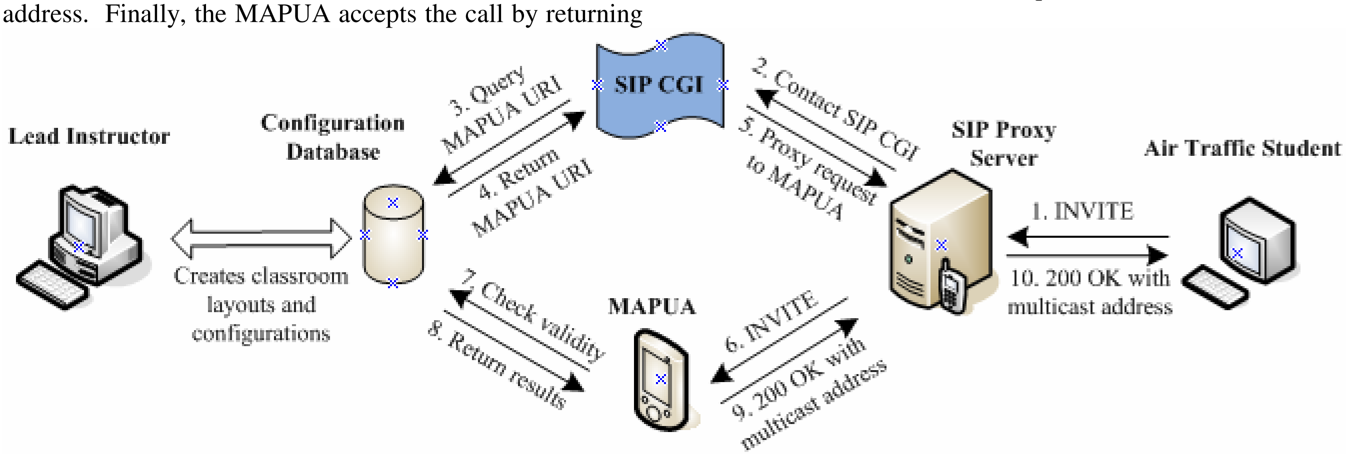

Figure 4. Call flow diagram for radio line communication

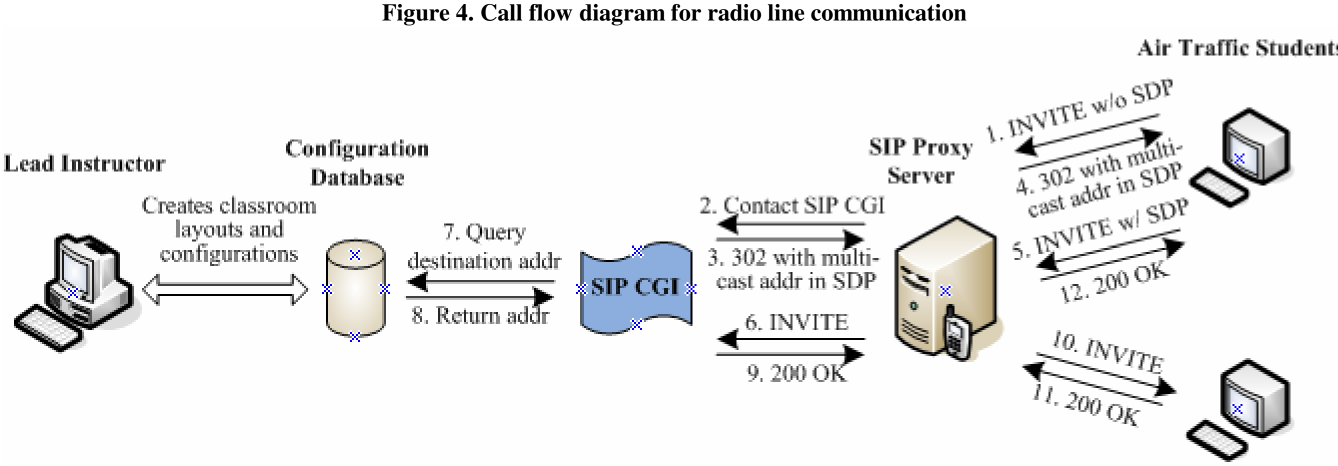

Figure 5. Call flow diagram for land line communication

The call flow diagram for land-line communication is shown in Figure 5. All the land-line communications have two distinct phases in signaling; first, where a unique multicast address is assigned and second, where the call destination is determined. When a student UA calls a land-line, the SIP proxy server executes the call logic associated with it, which dynamically assigns a unique multicast address and sends it back to the caller, requesting him to use that for the subsequent land-line communication. The student UA in turn sends a fully specified call request to the proxy again, which determines the possible destinations for that call based on the details of the caller and the type of the line, and forks the request to the destinations. The destination could be one or more stations, for example when an air traffic controller calls on a land-line, all the pilots in his sector get the intimation but only one pilot can acknowledge and accept the call.

3.2.3 Classroom recording

Classroom recording requires mechanisms to start and stop recording at the granularity of exercises, dynamically add and remove sessions during the course of recording, as well as mechanisms to browse through the recorded streams and play them back in VCR-fashion. All the record and playback functionality exists at the lead instructor UA and is accomplished using a Real Time Streaming Protocol [10] based design, in conjunction with SIP event notification framework [11, 12, 13].

The call flow diagram for classroom recording is shown in Figure 6. The RTSP client which runs at the lead instructor has to get the information about all the on-going communications. To accomplish this on a continuous basis, as the students enter and leave communications sessions, the instructor UA sends SUBSCRIBE requests to the student UAs. Whenever there is any change in their communication status, the student UAs NOTIFY the instructor UA. Once the instructor UA gets the multicast addresses of all the current sessions, the RTSP client requests the RTSP server to start recording. RTSP client can issue subsequent requests to the RTSP server, for adding (RTSP RECORD) or removing (RTSP TERMINATE) multicast sessions. The RTSP server captures media from all the required multicast sessions, mixes them and saves the unified media stream with the help of the file server. The lead instructor uses the file server interface to browse through the recorded files and play them in VCR-fashion using our archive player.

3.2.4 Instructor broadcasting and monitoring

Instructor broadcasting is designed similar to a land-line call, for which the destination is the entire classroom and which has a call logic that automatically accepts the call at all the receivers without the need for any explicit visual cue or acknowledgement. Such an automatic enrollment arrangement serves as a localized dynamic broadcast channel.

For instructor monitoring, the instructor UA does not need to perform any explicit signaling, as all the channels have been setup as multicast sessions. Rather, the instructor UA has to coordinate with the UA of the student to be monitored, to get a continuous update of its communication status and this is accomplished using the SIP event notification framework.

3.2.5 Notification of transmissions

Real-time feedback of the channel status uses the SIP event notification framework, which provides a framework for the student UAs, those involved in the communication and those seeking updates, to exchange status information in real-time.

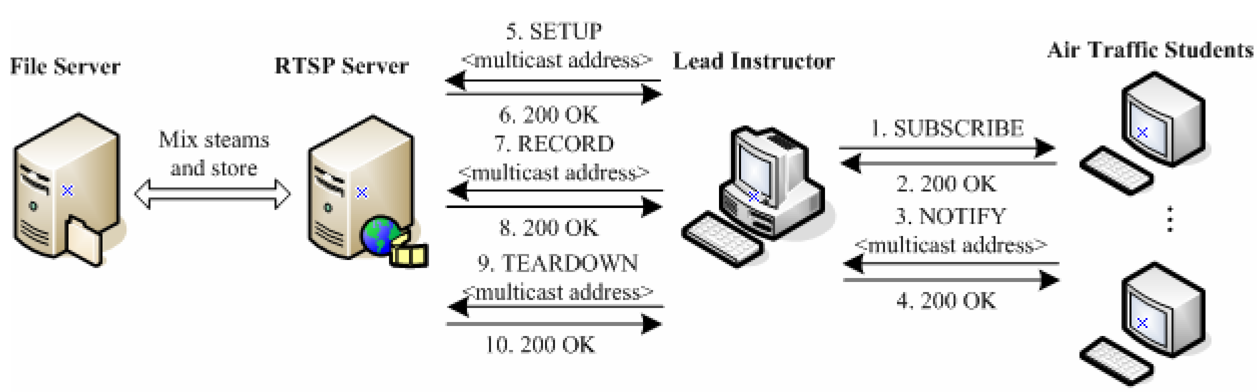

Figure 6. Call flow diagram for classroom recording

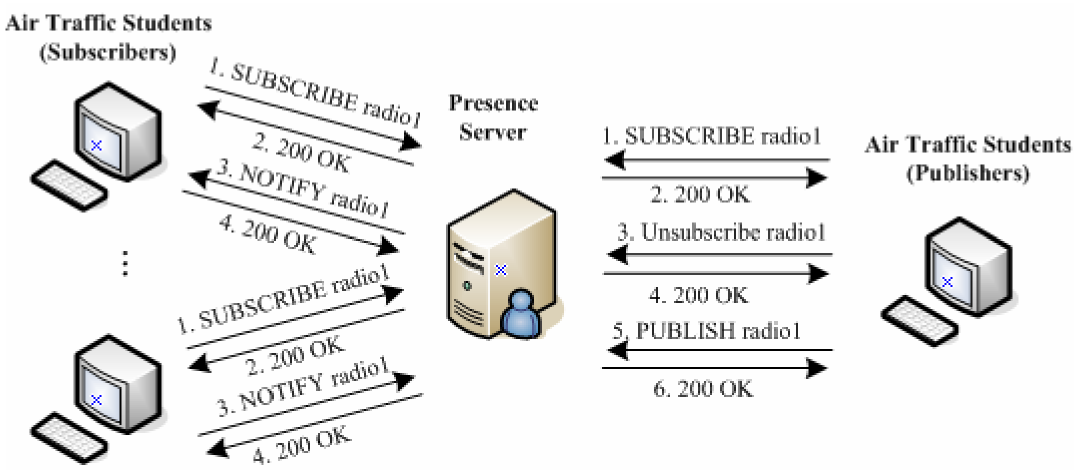

Figure 7. Call flow diagram for notification of transmissions

The call flow diagram for notification of transmissions is shown in Figure 7. When the training starts, no channels would be active and every student subscribes to feedback on all the channels. Whenever any student UA gets ready to communicate (either transmit or receive on a channel), it first unsubscribes from the channel updates, for it would know the channel status by itself. If a student UA starts transmitting, then it sends a PUBLISH request to the Presence server, so that NOTIFY messages are sent to everyone not on that channel but subscribed to its status.

3.3 Implementation

Since the research and development facility was away from the FAA Academy, we had to follow iterative prototyping as the system development methodology, with operation tests being conducted at the FAA Academy. Our language of choice for implementation was Tcl/Tk [14], a powerful scripting language well-suited for such rapid prototyping, with its platform independence and rich UI development tools. The file server, the archive player, the MAP user agent and all of the student and instructor applications including their user interfaces and user agents were implemented using Tcl/Tk. The user agents were built on top of Columbia's generic SIP user agent, SIPc [15]. All these aspects were made transparent to the end users by bundling all the required Tcl/Tk scripts into a single Windows executable application using Freewrap [16].

The media engine that runs on every student workstation and is responsible for all the media encoding and transfer was implemented in C++ with support for G.711, Speex, Internet Low Bit-rate Codec (iLBC) and GSM codecs. We leveraged CINEMA [17], an umbrella of SIP-based multimedia servers developed at Columbia, for building most of our VoIP server components namely the SIP proxy server (sipd), the RTSP server (rtspd), SIP-CGI framework, and general account management. Including only the relevant components of the CINEMA server, the project comprises of 70,000 lines of code. Finally, for the ease of deployment, all the deliverables are packaged as RPMs for Linux distributions and as Microsoft installers for Windows platforms.

4. EVALUATION

In this section, we discuss the important evaluation criteria including audio latency, packet loss and system reliability. The most common deployment scenario in the FAA Academy is to setup the student and instructor applications on Windows XP desktops and the VoIP server components on a RedHat Enterprise Linux server. All the machines are powered by Intel Core2 Duo 1.66 GHz or Pentium IV 3 GHz with 1 GB RAM, and connected to the Gigabit Ethernet backbone of the academy. Some advanced training rooms are operated on closed LANs.

To perform audio latency measurements, we setup the machines as shown in Figure 8. Two air traffic entities, namely an ATC student and a Pilot, have established a bi-directional communication channel. The audio generator output is split into two streams such that one of them goes directly to the left channel of the stereo recorder, while the other goes to the ATC student, which then multicasts it on the network to be picked by the Pilot. The Pilot then plays it out to the right channel of the stereo recorder. Thus, by calculating the time difference between left channel and right channel inputs at the stereo recorder, we can compute the mouth-to-ear delay. We used Cool Edit Pro [18] and PSEQ [19] to perform these measurements. The results showed that the audio latency was always less than 100 ms, for two machines connected to the same backbone Gigabit Ethernet infrastructure. Also, monitoring the network, we observed a zero packet loss in the operating environment.

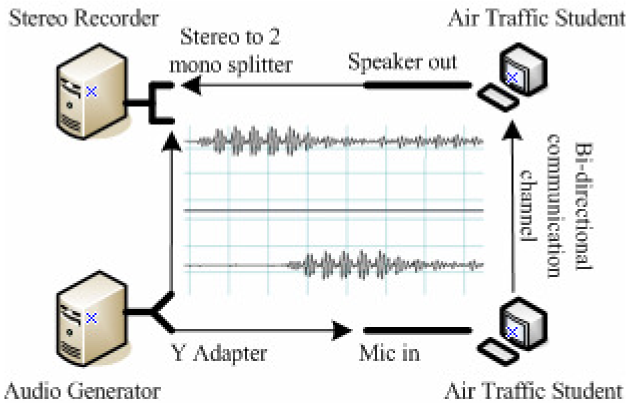

Figure 8. Setup for audio latency testing

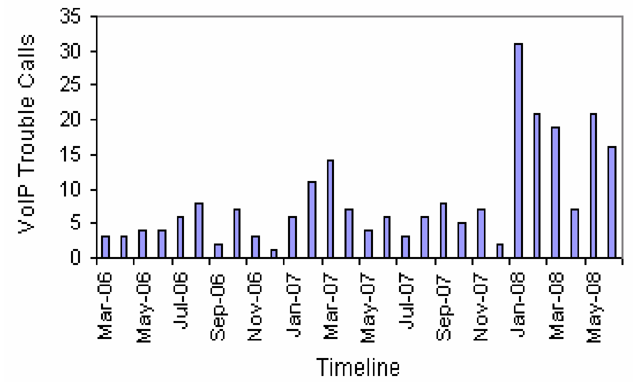

Figure 9. VoIP trouble call distribution

To measure the system reliability, we collected and analyzed all the VoIP-related trouble calls handled by the FAA Academy's information technology support staff in the five training rooms over the last 2 years, as shown in the Figure 9. This data gives an average of 1.8 VoIP trouble calls per room per month. These include a broad range of issues ranging from trivial ones like *VoIP system not started on a workstation* to more challenging ones like *the ESD problem in PTT* (discussed in Section 5.3).

The graph indicates two peak periods of trouble calls. The one during February - March 2007 was primarily due to a change in the instructional and technical support staff in two training rooms, with its associated learning curve for the new staff. The second peak in 2008 was primarily due to the ESD problem in PTT. Both these issues are discussed in detail in the next section.

5. DEPLOYMENT EXPERIENCES

With the system being deployed and operational in five FAA Academy training rooms since early 2006 and with two more being rolled out currently, the VoIP communication system has been a success. Handling of the project from thought to finish, from design and development to deploy-ment and user training, by Columbia University helped elevate this project from an academic research prototype to a production system in a federal environment. Equally appreciable is the openness with which the air traffic instructors and the academy system administrators embraced the new techno- logy. But all said and done, the deployment experiences have been full of ups and downs. In the following sections, we outline some of the significant operational challenges and reflect on the remedial measures. We group these challenges into three broad categories.

5.1 Conflicting mindset of the stakeholders

No one involved in the product lifecycle wanted it to fail but the most prevalent challenge has been to balance the conflicting mindsets of the involved stakeholders, when what was intended as a research prototype got deployed as a production system. The researchers have a dislike for anything that is not research; bug fixes do not lead to publications. The academy system administrators, who now had to support a new communication system, were not as confident and fluent about its functionalities and nuances as they were with the old system. The air traffic students, despite all their good intent, may turn unfriendly as any unexpected behavior of the training system would lead to lengthy class breaks. It is quite surprising that so far, the VoIP system has survived and continued to grow steadfast.

Continued interactions and collaborations amongst the stakeholders, from the early stages of the project, made sure that the problem of conflicting mindsets was at least well acknowledged, if not completely solved. We have shown two-fold commitments – first, by visiting the FAA Oklahoma facility six times in the last three years, to better understand the usage patterns and to identify and fix deployment issues, and second, by organizing 4-days extensive hands-on VoIP training sessions for the FAA Academy system administrators, thrice in the last two years. The FAA Academy on its part have shown good product ownership qualities – with the system administrators *hand-holding* the training sessions when the instructors are not fully familiar with the training systems and successfully resolving minor and repetitive issues locally and with the managers expressing their long-term commitment in the form of continued project funding.

5.2 Component Failures

Interfacing with a number of hardware I/O devices and software components brings new design and operational challenges to the communication system, where the system needs to continue functioning even in the event of malfunctioning of the associated components.

Some components become non-responsive if not being used frequently. For example an open connection handle gets automatically closed by the MySQL server after a preset period of client inactivity. Explicit care has to be taken while developing device wrappers so that all the critical events are captured and processed. For example if a PTT is inadvertently pulled out and immediately pushed in, the OS framework generates a different handler for that device. If this is not recognized and handled correctly, the applications continue to use an invalid device resulting in no I/O.

Self-correcting design: This led us to designing the component interfaces in a self-correcting fashion. By frequently storing the working snapshot of a given component, it can be restarted and its state be restored, whenever the main application detects a problem with it. For example achieving a seamless plug-and-play support for I/O devices by renewing the device handler upon detecting any hardware pullout/plugin events, or reestablishing the connection to the MySQL server whenever a "connection failed" error is received.

Faulty hardware: No design approach would be helpful if the hardware were to become physically faulty. A problem since the early deployment days has been the effect of Electro Static Discharge (ESD) on the PTT. For user convenience, the PTTs are designed with a metal clip that can fix the PTT base to the garment, while the student is on the move. This metal clip unfortunately serves as a good carrier for the ESD, when a student is charged from clothing or humidity level in the environment and has not been cautious to use ESD mats, thus spoiling the PTT circuit board. Such issues are not only difficult for the VoIP system to detect but also for the FAA Academy to fix – seeking a redesign of the PTT with the manufacturer seems as tough as asking the students to use the PTTs without metal clips.

5.3 Fault Diagnosis

Lost in Translation: With development and deployment environments being geographically separated, fault diagnosis was expected to be a tricky issue. On one hand, replicating the FAA Academy's training room setup at Columbia University was impractical due to expensive air traffic equipments and lack of skilled testers. On the other hand, the IT support staff working in a loaded schedule that has demanding uptime requirements, cannot always be expected to diagnose the faults precisely; the most commonly reported problem would read "*workstation 23 can not hear anything*."

The problem was especially evident in a prolonged trouble episode that lasted for about 2 months in early 2007, when the VoIP system was deployed in new training rooms to be managed by self-trained technical support staff. On occasions when a particular workstation experienced a problem with media transmission or reception (either due to inadvertent PTT pullout or the ESD problem with the PTTs, neither of which were known till that point), the VoIP application on all the training room workstations were being restarted – potentially, disrupting the training and worse yet, requiring the instructor to restart the exercises on many occasions. Since the PTT-related issues were neither detected by the logging mechanism nor clearly understood by any of us till then, it led to a situation where the researchers attributed this to incorrect administration and the system administrators took it for an incorrect system behavior. Only after a site visit by us, were the issues with PTT interfacing detected and heartbeat mechanisms were integrated into the system, along with options to restart VoIP system on selected workstations.

Remote Debugging: What complicates the issue further is that since the FAA Academy operates in a highly-regulated federal environment, security policy prevents any form of remote connectivity to any of their computing systems. Thus, we had to build a mechanism such that all the interactions of the system with the users, the network and the associated hardware had to be extensively logged. Reconstructing chain of events that leads to a failure, using log analysis is a painfully long process, with no guaranteed success – as not everything that needs logging would be logged and such experiences led us to rely heavily on self-correcting design, as described in the Section 5.2. Despite the best of the practices, there have been occasions where we had to fly down to the Oklahoma facility for live fault diagnosis.

Multicast support in network elements: For its correct functionality, our system expects that multicast be supported by the network elements like switches and routers. It becomes problematic when a switch has been designed to do something "intelligent" like multicast filtering, but with no control for the administrator to disable it. We ran into such an issue with a particular model of 3Com switch, only to be diagnosed after two weeks of remote debugging every other component in the system. Now our array of debug tools includes Mcast [20].

6. RELATED WORK

Not much information about air traffic training systems is publicly available. It is our understanding that MITRE Corporation Center for Advanced Aviation System Development [21] and William Hughes Technical Center [22] develop most, if not all, of the air traffic training software deployed at the FAA Academy. We are not aware of any other attempts to use VoIP in the air traffic training domain.

7. CONCLUSION

This research and development endeavor would not have been possible, if the Interactive Instructional Delivery System (IIDS) group of the FAA Academy had not looked for sustainable alternatives beyond the legacy system. This work is a novel case-study of a successful application of VoIP technologies beyond the Internet telephony. We also shared our experiences in deploying and maintaining the VoIP system over the last two years at the FAA Academy.

Our ultimate end-goal has been to transfer the technology and the product development responsibilities to a fully trained communication system staff at the FAA Academy – so that they not only take care of the day-to-day administration and management of the VoIP system but also take over the long-term product evolution, with minimal consultation from us. We recognize that this process is not going to be easy due to a variety of technical and non-technical reasons.

VoIP has been identified as one of the voice technologies for the proposed Next Generation Air-space System (NAS) [23] to be designed and deployed by 2020. It is natural to think about the applicability of our VoIP-enabled training system for real-world air traffic management. While the nature of the communication channels and the air traffic devices remain almost the same, the underlying network setup invalidates our design assumptions on security and scalability, thus making the training system completely unusable in the field as-is. It is our hope that the success of this training system, combined with existing large-scale deployment of VoIP in the real-world serves as stepping stones in this direction.

8. ACKNOWLEDGEMENTS

The project has benefited by the research and development contributions of Prakash Gowri Shankar, Venkata Sastry Malladi and Anurag Chakravarthi during the first two years of the project. On the FAA Academy side, the project was conceptualized, initiated and directed by Susan Buriak of the IIDS group. Michael Pierce of the Enterprise Services Center, Alice Ah-lo of ATCS and Martin Thee of the FAA Academy have helped us understand the air traffic training domain throughout the project. Our special thanks go to the administrative and instructional staff at the FAA for their continued usage, feedback and support for the VoIP system. This work is supported by multiple grants from the FAA Academy over the last four years.

9. REFERENCES

[1] Inspector General's report on FAA training of air controllers, Associated Press, June 2008, http://ap.google.com/article/ALeqM5gR4vVByQBFK39PU_3HbVj8jdxY4AD917H6PO3

[2] Key Issues Facing the Federal Aviation Administration's Controller Workforce, Statement of the Inspector General, U.S. DOT, June 2008, http://www.oig.dot.gov/item.jsp?id=2310

[3] ELO LCD touch monitors, http://www.elotouch.com/Products/LCDs/default.asp

[4] GVision LCD touch monitors, http://www.gvision-usa.com/touchscreens.html

[5] Plantronics air traffic controller PTT and headsets, http://www.plantronics.com/north_america/en_US/products/specials/air-traffic-control-headsets

[6] Kinesis programmable USB foot switches, http://www.kinesis-ergo.com/fs-savant-elite.htm

[7] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: Session Initiation Protocol," RFC 3261, Internet Engineering Task Force, May 2002.

[8] J. Lennox, H. Schulzrinne, J. Rosenberg, "Common Gateway Interface for SIP," RFC 3050, Internet Engineering Task Force, January 2001.

[9] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," RFC 3550, Internet Engineering Task Force, July 2003.

[10] H. Schulzrinne, A. Rao, R. Lanphier, "Real Time Streaming Protocol (RTSP)," RFC 2326, Internet Engineering Task Force, April 1998.

[11] A. Roach, "Session Initiation Protocol (SIP)-Specific Event Notification," RFC 3265, Internet Engineering Task Force, June 2002.

[12] J. Rosenberg, "A Presence Event Package for the Session Initiation Protocol (SIP)," RFC 3856, Internet Engineering Task Force, August 2004.

[13] A. Niemi (Ed.), "Session Initiation Protocol (SIP) Extension for Event State Publication," RFC 3903, Internet Engineering Task Force, October 2004.

[14] Tcl/Tk, http://www.tcl.tk

[15] Xiaotao Wu, Henning Schulzrinne, "SIPc, a multi-function SIP user agent", Proceedings of 7th IFIP/IEEE International Conference on Management of Multimedia Networks and Services, pp. 269-281, Oct. 2004.

[16] FreeWrap, http://freewrap.sourceforge.net

[17] K. Singh, W. Jiang, J. Lennox, S. Narayanan, H. Schulzrinne, "CINEMA: Columbia InterNet Extensible Multimedia Architecture", Columbia University Technical Report, CUCS-011-02, May 2002.

[18] Cool Edit Pro (now Adobe Audition), http://www.adobe.com/products/audition

[19] PESQ, http://www.pesq.org

[20] Mcast multicast traffic receiver, http://www.st.rim.or.jp/~yumo/pub/mcast.html

[21] MITRE Corporation Center for Advanced Aviation System Development (CAASD), http://www.mitre.org/about/ffrdcs/caasd.html

[22] William J Hughes Technical Center, http://www.faa.gov/about/office_org/headquarters_offices/ato/tc/tc_overview/

[23] NVS: Voice Switch for Next Gen Air Transportation System, http://www.faa.gov/about/office_org/headquarters_offices/ato/publications/oep/version1/reference/nvs/