

Leading-Edge Voice Communications for the MITC

Ben Teitelbaum <ben@internet2.edu >

September 12, 2003

Introduction

The Michigan Internet Technology Center (MITC) building represents a unique opportunity to showcase a leading-edge telecommunications system, demonstrate how real-time business communications may be conducted over a high-performance converged IP network, and engage staff in exploring new models of real-time personal communication. Voice over IP (VoIP) and integrated communications, which we define as *voice, instant messaging, video and other media in the context of presence*, have the potential to enrich personal communications, increase worker productivity, and reduce the recurring operational costs of business telephony. As the IETF standards for session signaling, instant messaging (IM), and presence mature, personal communications applications will emerge that unite voice, video, IM, file-sharing, calendaring, and whiteboarding under a framework of standards-based interoperability and automatic, rich presence. This white paper presents a vision for positioning the MITC building as a living laboratory for these new applications.



Figure 1: MITC Building—Draft Exterior

The communications systems and services deployed in the MITC building must be *leading-edge*, but not *bleeding-edge*. This building will serve as the

DRAFT

workplace and nerve center for some of the nation's foremost computer networking professionals, many of whom are technology leaders tasked with advancing the technology and use of collaborative communications applications in higher education. The system must also support the needs of non-technical staff, who simply need to perform their jobs. The MITC's core telecommunications systems should lead and inspire. They must also work.

From opening day in Fall 2004, the MITC building's communications system should offer new value and inspire new thinking about communications, without sacrificing reliability or requiring radical adjustments to employee work flow. The new system should not require substantial retraining or immediate readjustment to a new work flow paradigm, but should provide a clear opportunity for growth and experimentation, both at a personal level and at a system level. At a personal level, users should notice immediately that the MITC system is different from what they are used to, but that they also may interact with it in a conventional manner. At a system level, the design should allow new VoIP and integrated communications technologies to be embraced and deployed as they mature. Evidently, some tradeoffs *will* need to be made. These are discussed below, especially in the section titled "Systems Administrator Perspective".

Voice is the primary means of real-time communication. When managers assess status and give direction, when groups collaborate to reach decisions, when two parties negotiate to reach agreement, voice is usually the medium of choice. It is efficient, direct, and infinitely flexible. Most importantly, voice supports the informal, social communication (i.e. "small talk") that lubricates all business and professional collaborations. The importance and, in some cases, centrality and criticality of voice and electronic voice communications in the workplace is not to be debated. Voice communications systems must provide reliable voice connectivity *and* must function well for even the most naïve user. Because voice communications is *so important*, there is significant potential for new value in voice-centric systems that extend voice beyond what is offered by traditional telephony.

To address these competing goals, we propose a system that provides both traditional telephony reliability and functionality *and* advanced new voice communications capabilities. To accomplish this, user expectations will be managed carefully to create one set of expectations for traditional voice services and another for advanced communications features. The message to users will be this...

If it looks like an old-fashioned phone call, it is.
If looks like something new-fangled, it is.

Figure 2: A Dual System Requires a Dual Set of User Expectations



Figure 3: Example IP Desk Phones

The remainder of this memorandum describes the attributes of a system that could realistically be deployed by opening day (Fall 2004). This is followed by a description of advanced functionality that is not realistic for Fall 2004, but towards which we should nevertheless keep an evolution path open. Finally, I identify some open questions and suggest several next steps.

Opening Day, Fall 2004

This section describes a vision of a system that would be operational on opening day. It is described first from the perspective of a *typical user* and then from the “under the hood” perspective of a systems administrator.

User Perspective

Equipment

On opening day, the user will find a desk phone sporting a conventional handset, high-quality speakerphone, numeric dialpad, and the ability to contact emergency services by dialing 911. The telephone will support numeric dialing and basic PBX-style call control (hold, transfer, conference, and call forwarding, etc.). The desk phone will, however, be an IP appliance that is integrated with software running on the user’s PC to support advanced communications applications. Several examples of such phones are shown in Figure 3.

When the desk phone is used as a conventional PBX extension, users will experience a degree of reliability and functionality comparable to the PSTN. However, users will be told that when advanced features are used, normal expectations about voice communications should be relaxed.

Some advanced communications features may be supported directly on the desk phone. For example, these devices may sport a multi-line displays with the ability to receive text messages or perform directory lookups. Depending on cost, some or all phones may also support desktop video conferencing. Video output may appear either on the phone itself or on the user’s work-

station in an integrated PC-based application.

Most advanced features will be implemented in software that is run on the user's workstation or laptop. These features will be carefully integrated with the IP desk phone. The PC may serve to integrate additional media into a voice call, such as instant messaging, video, or whiteboarding. Or, the superior graphical user interface (GUI) and text entry capabilities of the PC may serve as control interfaces to voice-only calls; this will likely be the preferred way to initiate calls to non-numeric voice addresses and to control multi-party conferences. In all cases, the IP desk phone should serve as the primary audio device. Users will find the audio interface of the IP desk phone familiar and, in many ways, superior to what is available on the PC even in the context of an advanced communications session.

Advanced Features

On opening day, the user will be provided with the following advanced communications features:

- Retrieval of voice mail messages as email (i.e. “unified messaging”)¹
- Call forking (i.e. causing a single incoming call to ring multiple devices)
- Follow-me (i.e. allowing users to move from device to device without changing their “number”)
- An web-based company directory supporting “click-to-dial”
- A convenient interface for initiation of calls to textual SIP addresses (e.g. `sip:jdb@mit.edu`)
- A web helper application supporting “click-to-dial” for any SIP addresses or phone number (i.e. `tel:` URL) found on a web page
- Encrypted instant messaging
- Basic presence
- One or more supported “soft clients” to provide voice, IM, and presence on the laptop²
- Seamless multi-party IP voice conferencing

¹Voice retrieval of email is regarded as less important, though commonly a feature of “unified messaging” systems.

²This is especially important for remote or traveling employees.

The user who expects no more and no less of an office phone system than what is customary on a PBX, will be able to use the MITC system more or less exactly as if it were a traditional PBX. It is important, however, to send a signal to even the most conservative user that this phone system is technically exceptional. This is especially important for external callers calling into the system.

Identity

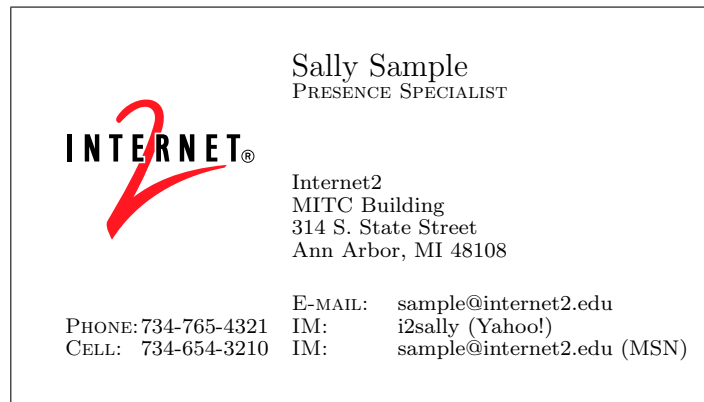
For reasons that are both technical and sociological, moving beyond E.164 telephone numbers is essential. Technically, decoupling the addresses of people from the addresses of devices, enables the engineering of a rich set of call processing and routing features. Sociologically, this decoupling sends a strong message to users that *something is different* and that they are empowered to help shape the future of interpersonal communications.

Each user will be reachable by voice not at a set of devices addressed numerically (e.g. +1-734-765-4321), but rather at a single user-specific address identical to the user's email address (e.g. `sample@internet2.edu`). This identity will serve as a general-purpose electronic identity good not only for email and voice, but also for video, instant messaging, presence, and other media.

Figure 4 illustrates how this convergence of identities might influence the text of a typical business card or email signature. In practice, there might be in interim form that presents the electronic identity `sample@internet2.edu` with explicit indication that it may be used for voice, email, im, and presence communications. In time, we expect that this will be implied, just as the “`http://`” has become implicit for most published URLs, and business cards will come resemble the “After” form shown in Figure 4.

Although users will be able to place external calls to E.164 numbers in the conventional manner, MITC desk phones will have no visible extension numbers, either internally or externally. Instead, each incoming call will be placed to either a user name (`sip:patti@internet2.edu`), a canonical alias (`patti.hogue@internet2.edu`), a rôle (`reception@internet2.edu`), or a location (`supply.room.aa@internet2.edu`). No per-user phone numbers will be assigned and new business cards will be printed to list the new address of the MITC building, the user's email/SIP address, and the phone number of an “Internet2 switchboard” (the same number for all users). Caller ID for calls outbound to the PSTN will show the switchboard number.

The “switchboard” will serve as a automatic call distribution (ACD) gateway between external callers calling from the PSTN and internal users with only SIP identities. As the external interface to the Internet2 telephone system, this ACD system serves a critical public relations function and should be designed with this in mind. Callers will be prompted by a voice recognition



Before

.....

After



Figure 4: Business Cards Before and After

system to “say the name or user name of the person you wish to reach or press zero to reach a receptionist”³.

Systems Administrator Perspective

Some leading-edge features of the opening day system will be under the hood and not visible to the typical user. An overview of the proposed system architecture is shown in Figure 5. IP handsets and desktop communications applications connect to a high-performance switched LAN core that provides external connectivity through one or more commodity internet connections

³This is deliberately designed to be somewhat painful for legacy PSTN users. In the event that this social engineering experiment backfires, we *will* have dedicated DIDs rented and allocated to each phone. These will be tightly guarded and made available only as needed.

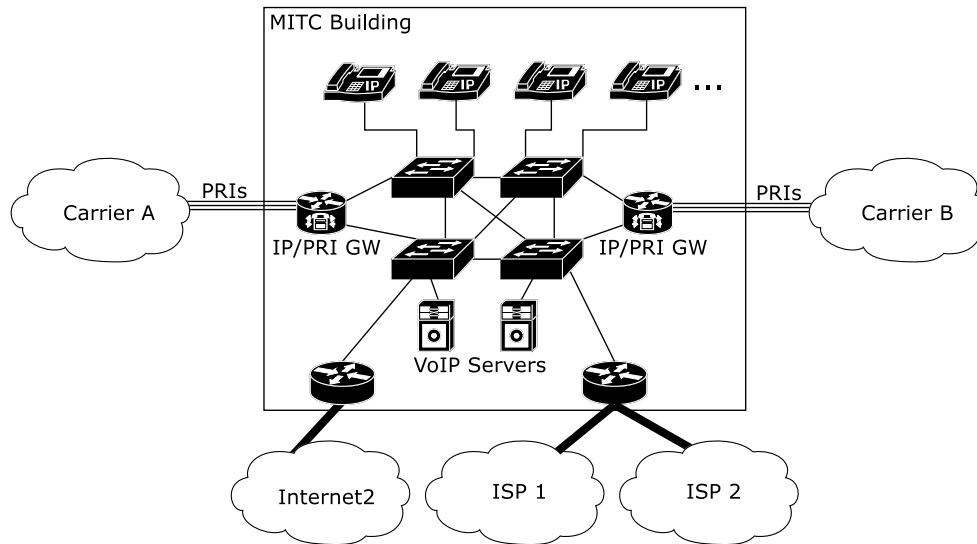


Figure 5: MITC Voice Network Overview

and an Internet2 connection (via Merit). These devices will also have connectivity to the PSTN via one or more IP/PRI gateways provisioned with voice trunks to separate voice service providers⁴. Finally, Figure 5 shows the “VoIP servers” that provide call signaling, routing, and voice mail functionality.

Although users will perceive that there are two ways to use the system, each with its own set of expectations, this duality is not carried through at all levels of the internal system architecture. On the server side, it is far simpler to build and operate a single system that supports different modes of use with different reliability profiles than it is to build and operate two systems. On the LAN side, engineering for complete duality would seem to preclude integrated use of the IP desk phone in advanced communications applications. For example, if the IP desk phones were given private addresses and connected via a dedicated LAN infrastructure only to the necessary servers and gateways, it would be impossible for a desktop communications application to signal the IP desk phone or for the IP desk phone to be used in end-to-end IP communications sessions (except those that are wholly within the building).

This conflict in the system design goals requires us to make careful engineering tradeoffs to achieve a leading—but not bleeding—edge system. One necessary tradeoff is that the system, even when used in its most conservative mode, will be subject to security vulnerabilities that are not present in conventional TDM voice systems.

⁴At least one set of TDM voice trunks should be brought physically to the MITC data center. Fallback gatewaying to the PSTN may be provided either by a second set of TDM voice trunks or through a separate provider with an off-site gateway.

The remainder of this section separately discusses various components of the internal system architecture (servers, wiring, and LAN engineering) and concludes with a discussion of security vulnerabilities.

Servers

A number of critical functions, including call signaling and routing, call control, conferencing, and voice mail, will be provided by “VoIP servers”. We expect these services to make heavy use of the emerging IETF standards for real-time signaling, messaging, and presence (SIP and SIMPLE). Multiple servers may be needed to provide various services and to improve redundancy. As with DNS, off-site secondaries should be used.

There are both proprietary and open-source SIP server options that are very strong. In either case, we require extensibility and interoperability to support future services.

Wiring

To meet the expectation *if it looks like an old-fashioned phone call, it is*, a highly conservative wiring plan and path to the PSTN will be provisioned. Each IP desk phone will be attached to a dedicated, voice-only Category 5 (or 6) Ethernet drop, supporting 802.3af power over Ethernet (POE). Calls between these phones and the PSTN will be transported over these dedicated drops through a redundant Ethernet switching fabric to the MITC data center and, from there, to one of several IP-PSTN gateways each connected to TDM trunks from separate local exchange providers.

IP handsets, core LAN switches, inline power inserters, VoIP servers, and IP/PRI gateways must all continue to operate in the event of a power failure. The MITC building is being engineered to provide critical systems with uninterrupted power through the use of an on-site backup generator. The voice infrastructure must be accounted for in these plans.

LAN Engineering

Reliability can be increased further by separating voice traffic in the switched Ethernet LAN core. A separate “voice LAN” (probably a VLAN) should be created for IP desk phones, as well as the system gateways and VoIP servers. To protect voice traffic from traffic-based denial of service (DoS) attacks, we recommend a QoS-enabled VLAN, access to which is policed both on the basis of MAC address and physical port. Each physical port on the voice VLAN will admit traffic from only a single permitted MAC address. This approach helps to protect the voice VLAN from attack, prevents accidental

damage (e.g. someone plugging a high-bandwidth video source into the Ethernet jack), and allows the PSAP to be provided with reliable in-building location information.

To support advanced communications applications, devices on the voice VLAN will be given public IP addresses⁵. Traffic between the voice VLAN and the rest of the LAN or between the voice VLAN and the WAN would pass through one of the routers in the MITC data center, where it would be policed carefully to protect voice traffic. Specifically, multicast traffic must be blocked from entering the voice VLAN (because many cheap IP phone devices take a CPU interrupt on every multicast packet and seize up in the presence of even small amounts of multicast traffic). Furthermore, a cap would be placed on the total aggregate of traffic admitted to the voice VLAN from the outside and caps would be placed on the amount of traffic that could be sent to any particular device on the voice VLAN.

Security Vulnerabilities

Although the precautions mentioned in the previous subsection protect the voice VLAN from traffic-based DoS attacks, they do nothing to protect against protocol-based attacks. Such attacks exploit protocol design or implementation defects and are unfortunately common⁶. Possible implications of protocol-based attacks include not only denial of service, but also theft of service (toll fraud), loss of privacy (unauthorized wiretapping), defacement (e.g. of the voice prompts on the voice mail server), and identity spoofing (e.g. the initiation of calls that appear to originate from Internet2, but do not).

To protect against such attacks we recommend robust testing of all IP phone devices acquired. The IP handsets are invariably appliances and must be purchased as such. VoIP servers, directories, and gateways usually run on common operating system platforms. Unfortunately, these are also often sold as “black box appliances”—mostly to simplify customer service for these companies. We recommend whenever possible to build and secure our own servers before installing and running any needed IP telephony server software. The vendors of such “server appliances” have a poor security track record and all claims about the security of such “black box appliances” should be suspect. We also recommend continuous performance monitoring and the deployment and operation of an intrusion detection system for the voice VLAN. *Implementing rigorous security practices such as these will require additional staffing resources.*

⁵This will increase Internet2 and Merit’s overall demand for IPv4 address space and must be planned for accordingly. IPv6 VoIP solutions will not be ready within the time frame of the initial deployment.

⁶See: <http://www.ee.oulu.fi/research/ouspg/protos/testing/c07/sip/index.html>.

Finally, it may make sense to police access to the voice VLAN by protocol and TCP/UDP port number, though this should be avoided if possible. Because the devices on the voice VLAN will be centrally purchased and managed, there will not be great heterogeneity. Hence, there is not a significant threat of there being unneeded open service ports available on the VLAN. Firewalling or NATing voice traffic is additionally complicated by the fact that SIP dynamically allocates UDP ports for media streams. This means that pinholes must be punched through any firewall on a dynamic basis as calls are initiated and received. Though there are tunneling protocols for accomplishing this (e.g. STUN and TURN), it adds significant complexity and fragility to the system, perhaps without improving significantly improving security.

Advanced Functionality (Beyond 2004)

- Seamless multi-party IP video conferencing
- Wideband audio (e.g. 16kHz sampling)
- Automated, “activity presence” and call forwarding derived from enterprise calendaring
- Audio presentation of presence from the “switchboard” ACD system. For example, in the event a person is unavailable, the ACD system might say: “Sally Sample is unavailable for a phone call for the next 36 minutes. You may press star to leave voice mail or send an instant message now to sample@internet2.edu”.
- Automated, spatial presence derived from DHCP, GPS, and/or Bluetooth location beacons; clients will learn not just their geographic location (longitude, latitude, altitude) and civic location (“MITC Building, Room 214; 780 S. State St; Ann Arbor, MI”), but also the location type (“office”, “conference room”, “lobby”, “home”), hints about the etiquette and privacy of the space (e.g. “public”, “private”, “quiet”), and the correct public safety access point (PSAP)
- Support for one or more wireless-friendly voice codecs, including a media gateway to translate to/from G.711
- Strong end-to-end encryption for both signaling and media
- Support for IPv6

Open Questions

- Which organizations within the MITC building will be served by the system? Internet2? Merit? Tenants of MITC (e.g. in the incubator space)? Conference center? Data center? Other building tenants?
- Will we overhaul telecommunications in Internet2's DC office as well? If so, would we try to work with EDUCAUSE? How would this be integrated? Would we need a QoS-enabled VLAN that spans both sites?
- What is the role of the building owner?
- Will there be a PBX in the building?
- To whom do we outsource the SIP-PSTN gateway functionality? Options include: the University of Michigan, an internet telephony service provider (ITSP), a carrier, or us (i.e. no outsourcing).
- To what extent can we partner with the University of Michigan to deploy and operate leading-edge functionality?
- What are the requirements for analog lines (e.g. Merit requires some for modem testing)?

Costs

It is important to understand even at this very early stage of planning what this system might cost.

TBD

Next Steps

1. Answer the open questions above
2. Work with the University of Michigan to understand how we can partner with them⁷
3. Determine the budget

⁷Some initial discussions with Pradip Patel have already occurred and have been encouraging. He is working closely with Nortel testing a highly-scalable SIP solution that could meet our basic needs and many of the "advanced" requirements set forth in this document.

4. Engage potential solution providers in exploratory discussions on their product offerings, technology directions, and interest in corporate sponsorship

Glossary

ACD Automatic Call Distribution; an automatic system that duplicates the functionality of a traditional receptionist with respect to incoming calls

Black Phone Non-IP telephone (analog or digital)

E164 The international public telecommunication numbering plan

E911 An two-phase FCC program to require wireless carriers to provide PSAPs with the physical location of a 911 caller

Hard Phone Physical IP telephone

Integrated Communications Voice, instant messaging, video and other media in the context of presence

PSAP Public Safety Answering Point; where 911 are routed

SIMPLE SIP for Instant Messaging and Presence Leveraging Extensions; IETF working group developing extensions to SIP to support instant messaging and presence

SIP Session Initiation Protocol; IETF signaling standard for initiating, terminating, redirecting, proxying, and performing other operations on real-time multimedia sessions

Soft Phone Software only implementation of an IP telephone

UA SIP User Agent; could be a phone (hard or soft) or a logical component of a middlebox

URI Uniform Resource Identifier; a compact string of characters for identifying an abstract or physical resource; *e.g.* `sip:bob@bigu.edu`

VoIP Voice over IP (Internet Protocol)