Draft Architecture and Roadmap for Communications Services at MIT

October 30, 2006

Motivation

A transformation of how we communicate in our professional, academic, and personal lives has already begun. But this area will become even more dynamic as new services and communications devices are developed and commercialized. More importantly, we will all discover new ways to utilize these communications tools in our everyday activities. Life on the Internet will become even more "virtual" and with the adoption of various communications models for the various groups and communities with which we associate.

Much of the volatility in this area is caused by the tension between traditional communications services and products and those being developed for the Internet. Traditional services such as those offered on the Public Switched Telephone Network (PSTN) are based on communications controlled by the network. In contrast the Internet has fostered innovation by moving control to the edges of the network and by putting that control into the hands of both large and small enterprises, and on individuals themselves. Legislation and regulation, both local, domestic, and international, will also impact the evolution.

Many of these changes will be driven by personal choice. Today, the mobile phone is an example of this choice. The device that you carry with you everywhere should meet your specific needs. But it must also be able to interoperate with multiple networks and a variety of services. As the number of available networks increases and as new services become available these choices will become more complicated. With the explosive growth of available products, more and more users will experiment with them and adopt them as part of their communications routines, raising expectations for Institute provided services.

Finally, there is an opportunity for the Institute to define a set of services and build an infrastructure that ties this all together into a coherent communications environment that is both highly functional and yet adaptable to the needs of individuals and the various communities with which they communicate. This environment will be based on a variety of fixed and mobile devices, networks, and services. And there will also be the need to allow our users to seamlessly use services provided my the Institute, by other academic and research

organizations, and by other personal communications services of their choice. A sound architectural model and a shared vision for how services may evolve will lead to a effective roadmap for the logical and efficient deployment of services at the Institute.

Scope and Goals

The purpose of this document is to lay out a draft architectural model and evolving plan for the advancement of current services and the introduction of new services. The communications services area broadly includes all services that facilitate non-face-to-face interactions. This includes both real time communications, near real time communications, and asynchronous communications services. Voice, video, display (document, application, screen sharing), and instant messaging (IM) media are all included. The ability to attach and transfer documents in additional formats and to record and store communications sessions are also included in the scope.

This initial draft is intended as a "first cut" in order to solicit feedback and generate discussion. Hopefully this will identify items such as missing components, missing opportunities, clarity, and areas in need of more analysis. This document will evolve both based on this feedback and on changes in the environment surrounding these services.

The draft architecture will provide a reference model to facilitate these discussions and it will also expand and evolve over time.

The roadmap will both paint a longer term vision that can be shared with technology and business leaders at the Institute as well as provide direction for Information Services and Technology's planning efforts. The goal is to publish the roadmap periodically so it can act as a planning tool for shorter term activities including the following.

- Evaluation of emerging technologies and products
- Survey of standards evolutions
- Assessing changes in Institute and community requirements for communications services
- Defining and launching projects to deploy new services or add major functionality to old ones
- Defining and launching projects to retire old services and/or to migrate users to their replacements

- Identifying improvements to infrastructure components in other areas or defining requirements for new components that can be used to support communications services

Methodology

This communications services roadmap will first identify set of guiding principles for consideration in evaluating options for moving forward. A set of driving factors will also be outlined. Finally a set of architectural components will be identified with a vision for the future in each area. For each of these components a set of actions to be considered for both the short and longer term will be identified.

This roadmap will then be distributed more widely within IS&T, the Institute, as well as with vendors and colleagues at other educational and research institutions. Feedback from these discussions will be used to update and expand the next iteration of the roadmap.

Principles

The following set of principles is presented as a guide for building the roadmap. They broadly describe the desirable characteristics for communications tools and services. These principles are intended as the desired state and can be used to identify any compromises that need to be made due to the limitations of a specific product or technology, and at a particular point in their evolution.

Identities and addressing - Communications may be with other individuals, services, or devices. Addresses should be meaningful and have context. For example, user@mit.edu is preferable to 31234; helpdesk@mit.edu is preferable to +1-617-253-1101 (though this does have the 617 geographic area context). Devices such as traditional telephone sets, may only need to be identified by a phone number or MAC address.

Backwards compatibility - Even advanced communications tools should support "least common denominator" functionality to allow basic communications with older tools that have limited functionality. For example, a conference rooms multimedia collaboration system that supports video and wideband audio should still be able to communicate with a traditional analog telephone.

Ease of adoption, ease of use - New communications tools should be able used with a minimal amount of preparation, training, and documentation. It should be easy to "get started" with basic functionality with training and detailed

documentation only needed for more advance features. User support requirements should be evaluated.

Interoperability - All communications tools with common functionality should be able to interoperate. You should not have to have two devices at hand in order to do voice communications with any two groups of people or services. Gateways should be available to interconnect different technologies where needed.

Open federation - Communications services should interconnect with like services deployed at other institutions and corporations. They should also interconnect with public services offered on the Internet and the PSTN. Communications "islands" should be avoided.

Component services - Communications services that are build on open standards are highly desirable. The ability to interconnect service components allows for the construction of systems that are flexible and adaptable. The ability to replace components minimizes the possibility of "vendor lock-in".

Leveraging other systems - Communications systems should build upon existing components that can be reused or adapted for new functionality. Common systems used things like authentication, directory, databases should be leveraged whenever possible to maximize the value of existing systems.

Value - Existing and potential communications services should be evaluated based on the value that they provide to the Institute and it's faculty, staff, and students. Characteristics such as increased productivity, safety, mobility, flexibility, accessibility, reliability, and functionality should be weighed against the cost of deploying and supporting a service. Costs of supporting legacy systems should be considered in their retirement plans.

Driving Factors

The actions indicated by this roadmap are based on a variety of factors. And more important, they are expected to change over time and hopefully enable more options and choices. Part of this exercise is to identify "intersections" - the places where the Institute's roadmap will come together with other roadmaps such as those for standards or commercial products. It will also be important to identify future use cases where a users' communications needs will intersect with newly available technologies. Major forces include the following.

Standards - New standards for supporting Internet communications continue to evolve. SIP and XMPP are still relatively young protocols and new standards are being developed and refined. In some cases it may be wise to wait until the necessary standards are in place - or at least the direction is clear.

Commercial and open source products - Similarly, new products are appearing at a rapid pace. It is important to understand current product roadmaps and anticipate the appearance of new products. Products in this sense include both devices such as mobile phones and system components such as a rich presence agent.

Other services - It is also important to understand the evolution of the services provided by both peer organizations and those that are commercially or publicly available. There is always an expectation, or at least a desire, to provide interoperability with other services.

Increased connectivity and mobility - Broadband penetration is quite significant at this point but will continue to grow rapidly. Both wired and wireless connectivity are assumed on the campus. The majority of our users have broadband service for their homes. Hotels and hotspots provide access when we are on the road. And mobile and municipal networks will evolve to provide higher speed connectivity in between. All this will drive increased demands for services.

User expectations - Users will expect services that they can get elsewhere to also be available from MIT. Incoming students will expect that the functionality that has been important to them will continue to be available while they are here. Users will expect that they will continue to easily communicate with their old contacts and communities as they build new ones.

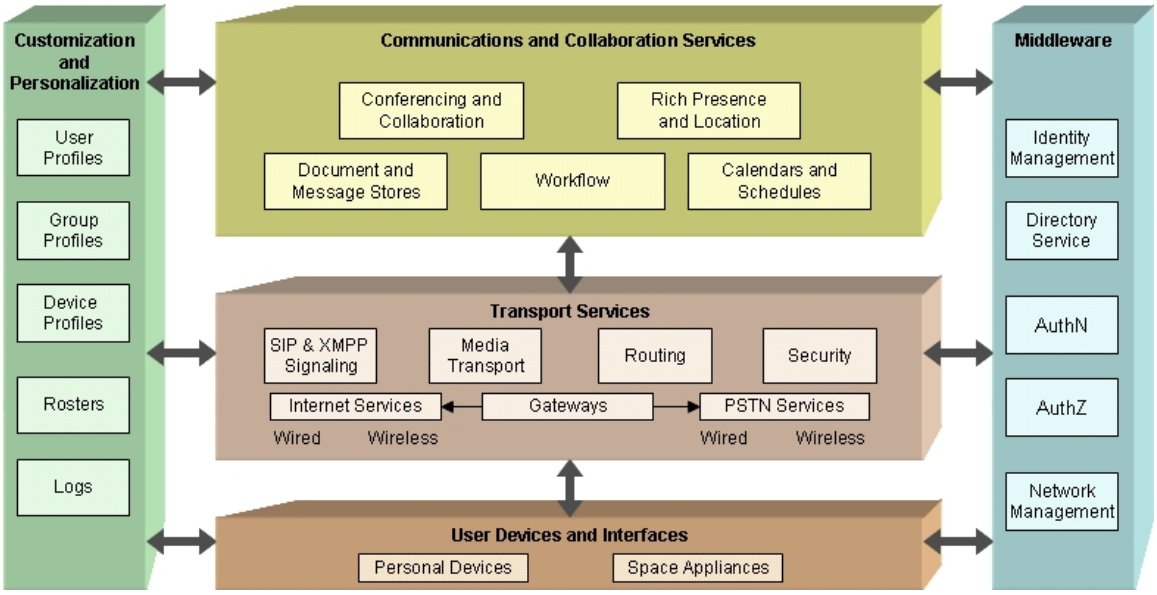
Anticipated needs - In addition to understanding the current expectations of our users it is important to be able to anticipate their future needs. Likewise, emerging needs of the Institute to support new activities and programs will need to be taken into consideration.

Communications Architecture Components

This section outlines the various components of a communications architecture and describes the characteristics, function, and importance of each. For each component a set of actions to be considered for both the short and longer term is identified.

The following diagram defines five structural component groups. Each of these groups have a varying degree of specificity to the communications architecture. While the user devices are seen as particularly aimed at communications services, those in the middleware group are components that should be reusable to support other services. The boxes within each group are included as examples of the functionality provided and not as a comprehensive list of components.

Communications Architecture





Devices

User devices are those that present a local interface to the user. They can be classified as mobile, personal devices that belong to the user and are carried with them, and also as communications appliances of various sorts that are stationary and are associated with particular spaces.

Vision

With the widespread access to the Internet on campus, at home, at hotspots, at hotels, etc. it is easy to be "connected" whenever you desire. The mobile phone network provides an additional set of features to allow even greater connectivity. And the trend to "always on" will continue to the point where you will have to make a conscious effort to "unplug".

A personal device will certainly be a key communications tool in the future. Today that device is typically a mobile phone – used primarily for voice but also for IM or email by some users. In the future such a device will continue to provide more features such as sensing the activities of the user, location, availability of networks, surrounding devices and even other people in the vicinity.

As mobile personal devices evolve spaces will also need to evolve, including a richer set appliances that will be available at fixed locations. For today's office a traditional desk telephone is that appliance. But why do you need a desk telephone if you can call and be called from the personal device in your pocket? There may be reasons such as quality of a speakerphone - but it is anticipated that other capabilities will emerge. An office of the future may have a wall with a large display, a set of surround sound speakers, a face tracking camera, etc. These devices should all be available to the person in the office and interoperate with their personal device. Unlike the traditional telephone all these devices will cooperate with each other and will be able to take on the identity of the user, be controlled by him, and act in his behalf.

Most Internet telephone devices today have limited security functionality. Most don't allow for the encryption of the media - encryption of the signaling channel with TLS is just becoming available in some models. In addition these devices do not have the capability to authenticating users - only the ability to authenticate themselves. If a user wants the device to take on their identity they have to enter a password into the device, where it is stored until deleted.

Future devices should have the ability to do actual user authentication. These devices may deliver a limited set of functionality for users that don't authenticate - for example a telephone might allow for emergency calling without authentication, providing a location to the emergency responder but not a persons name.

It will be difficult to enforce security mechanism until device capabilities evolve. It may be desirable to provide separate systems with lower and higher security until this evolution is complete.

Other spaces will similarly evolve. Conference rooms may have even more enhanced communications devices. They may be able to coordinate the communications with groups of people by doing things like raising a threshold for incoming calls while a meeting is in progress. For example, all but emergency calls could go to voicemail while allowing for IMs but setting the users presence to "in a meeting".

Spaces off campus will evolve in differing ways - particularly spaces where people gather or wish to communicate - the living room at home, the home office, the kitchen may all have different capabilities. The car will become a personal or

group communications space depending on occupancy. Other transportation systems will take on yet another set of capabilities. Where there were once "no smoking" areas there may "no voice or video" areas.

Actions

This area will evolve slowly over time. The mobile phone will probably continue gain new features but they may be more closely tied to the functionality specified by the mobile carriers than those desired by the user. We should continue to explore ways to provide better access to services such as email and voice mail using these devices. We should also explore the capabilities for mobile devices with Wireless LAN (WLAN) capabilities and, in particular, the ability to provide service transparently whether the device is connected to the Internet, mobile carrier network, or both. WLAN battery requirements and the desire of carriers to prefer the use of their networks may limit this functionality in the short term.

Software based clients are just emerging that that have the capability to support both the SIP and XMPP standards. We should evaluate these clients as they become available and consider selecting one or more as fully supported application.

Traditional telephones will probably continue to expected as a desktop tool for some time. However there is no reason not to continue to leverage the alternative of providing IP based phones based on SIP versus traditional PBX telephones. These phones will continue to evolve with additional features, improved voice quality, and enhanced security.

In the longer term there will be an opportunity for a richer communications experience as both mobile and fixed devices become aware of each other and gain the capability to cooperate with each other. At that point it will be worth reconsidering what appliances are installed in various spaces.



Transport

These components range from the physical copper wires and optical fibers that make up wired networks and the radio spectrum used for wireless networks to the packet and circuit switching networks that all carry communications services. Higher layer services will provide for the transport of signaling and media streams and the routing of messages used to create, modify, and terminate communication sessions. Various gateway functionality will be necessary to

translate between different signaling, media, and addressing schemes. Transport components need to incorporate security mechanisms to prevent penetration and disruption of services.

Vision

Communications services of the future will be built using the Internet Protocol (IP). This shift has already begun with the shift of investments from the traditional circuit switched PSTN hardware to IP based soft switches and packet networks. Mobile carriers are also providing advanced services based on IP and also offering pure IP connectivity. While the traditional PSTN and current mobile services will exist for a considerable length of time, the future will be in IP based solutions. During the transition gateways provide by the Institute or carriers will be necessary to interconnect the two network domains.

The reliability of IP networks will continue to improve. But as they mature so will the capabilities and motivation of those that want to use them for antisocial purposes. Spam for IP telephony (SPIT) and spam for IM (SPIM) will become problematic and mechanisms to control the impact will be necessary. The best solutions will make SPIT and SPIM not worthwhile to the perpetrators and the sooner that preventive measures are in place the more likely this will be achievable.

Actions

As part of this transition we should continue, and accelerate upgrades to the campus networks - both wired and wireless - to better support Internet-based communications services. Upgrades should include the following:

- Upgrades of network switches
- Deployment of VLANs to better support real time communications
- Increased redundancy of network components and links
- Expanding wireless coverage to the entire campus including outdoor areas

We are currently building communications services for voice and IM based the SIP and XMPP standards. We should continue to build on these standards with SIP being the foundation for voice and video and XMPP being the preferred standard for IM and presence services.

Video conferencing systems and services are still mainly based on the H.323 standard with systems that support both SIP and H.323 just becoming widely

available. Any new systems that are acquired should also support the SIP protocol for compatibility with our core communications infrastructure.

As standards are developed we should add the capability to authenticate SIP communications as a preventive measure against SPIT. Users should be able to choose how unauthenticated communications are handled. This will allow for the introduction of the concept of authenticated and unauthenticated communications before SPIT becomes widespread. This capability will only be effective when it is widely adopted but now is the time to take the lead and be prepared.



Services

Communications and collaboration services define a portfolio of applications that facilitate both individual and group communications. Some of these services may be highly specific to communications activities while others, such as document and message stores will be able to leverage the same technologies used in other areas. The services in the diagram are mainly shown as examples and new services should be introduced as new products and applications become available.

Vision

Users should be able to choose from a diverse set of applications, using those that are most appropriate for each communications activity. However, these services need to “hang together” in a cohesive way with clear functionality. similar interfaces, and shared common configuration options. Two examples are given here as core services.

Presence - Presence functionality can range in complexity from providing basic contact information - if and where an individual is currently reachable on the network - to sophisticated means of conveying the context currently surrounding the user and their immediate needs and desires. All of these express the ability and willingness of that individual to communicate to others that wish to do so. Presence information can also be applied more broadly to other identities. For example, services and devices may also make their availability known so others can contact them.

Basic presence information may consist of one or more network address where an identity may currently be contacted. This can enhanced by providing information on the types of devices that are reachable, their capabilities (media types, etc.), their geographic location, etc. Information about the devices' user

can also be presented including activity (eg. typing), who they communicating with (physically or electronically), or even their mood or desires. This rich presence information opens doors for applications that can act on your behalf to establish or schedule communications sessions with individuals and groups when they would be most effective. Finally presence information can be coordinated with other information such as a individual's personal calendar or physical location.

Presence information can be managed by a local client application or, in more complex systems, managed by a presence server. This server can compose complex presence information from various source including the user's devices, network sensors, and other services. Presence information can be provided to watchers by either the client (peer to peer) or by a presence server. Either can allow a user to control how much information is provided to a particular watcher, or class of watchers, and the granularity of that information. For example, you may choose to reveal that you are "online", or "at work", or "in your office", or even "talking to Bob about the budget".

Various vendor solutions currently implement presence services in some form or other. To allow for both intercampus interoperability and interoperability between various campus services and users, we should look for solutions that are based on standards such as SIMPLE or XMPP. Standards in this area are still evolving as are the mechanisms for generating and managing presence information.

Collaboration – Collaboration services must support a wide variety of media including voice, video, and text. Users also want the ability to share screens, specific windows, applications, and data files. Not only is there a need to record sessions but it is desirable to be able to edit and index the recordings for later reuse. And, as with other applications the collaboration tools need to interoperate with wide range of devices running different operating systems.

While providing complex functions these tools still need to be intuitive and easy to use. It should be simple to create sessions by going to a web page or by selecting an option in another tool such as a soft phone or IM client. The ability to add users, add media, and change the focus between users should be fluid. These tools should easily support one to one, one to many, and many to many sessions.

This extensive set of requirements leads to the large gap between the applications that are available today and what would really be usable and valuable to users. Many of these applications have clear roots in PSTN teleconferencing – with other media and screen sharing applications added. The evolution of these products has not let to the necessary revolutionary change that is needed for them to be effective in a diverse set of uses.

Actions

Both the current Zephyr and Jabber services provide basic IM and presence services. And many users not only use other services such as AOL Instant Messenger, Skype, and Yahoo! but use other Jabber services not provided by MIT. We should continue to encourage Jabber use at MIT and promote the use of the MIT provided identity. This will not only enhance the recognition of "user@MIT.EDU" as a single identity for all communications but also facilitate federation with Jabber services at other institutions, as well as Jabber based services such as Google Talk.

We should continue to support IM and presence clients that use XMPP but also support the use of multiple services and protocols. This will reduce the number of applications that user needs to run simultaneously. We should pursue clients that also support other media including voice and video, and also clients that support the SIP protocol in addition to XMPP. This can be the beginning of a ubiquitous communications client that will support both presence and multimedia communications.

A web based service that ties communications activities is also worth investigating. For example, presence and IM capabilities could be integrated with the Webmail service. This same interface could be used for managing voice mail messages and to initiate voice and video calls using third party call control – together providing a integrated communications interface via the web.

While reproducing all the functionality that has been provided by Zephyr is probably not necessary, we should enhance the Jabber service in an attempt to draw users to use it. Once we have reached a critical mass of MIT Jabber users we can begin to encourage users to transition from Zephyr.

Location information can be used in a variety of ways from the inclusion in published rich presence to providing the ability to automatically locate individuals contacting emergency service providers. For users directly on the MIT network we have the capability to locate them using current network management tools. We should add the capability to tie this location information into the SIP and XMPP services so that location information can initially be made available to user devices and to the Campus Police for emergency calls. Users should be able to have access to this information via the current web pages and have the ability to manually input location information for cases where it cannot be determined by the network. It is also anticipated that this information will be valuable to campus experiments such as iSPOTS and Living The Future (LTF).

There are various voice conferencing services currently available. We should evaluate the need for an MIT provided service that would allow for users to set up conferences "on the fly" using a simple web interface. The service would use the existing SIP and XMPP services for voice and IM. In the mean time we

should continue to look for emerging commercial or open source products that both fully support these standards meet the requirements for an easy to use, multimedia, multiplatform collaboration service.

The current SIP service provides for the blocking of Caller-ID information but this only provides a limited level of privacy on calls to the PSTN. We should also evaluate the need and alternatives for providing a true anonymous SIP service.

As these and other services are developed they should be integrated whenever possible with other services that facilitate effective collaboration. For example, the integration of presence and calendar services can tie together real time availability with future availability. The ability to store voice messages along with email messages and other documents in a flexible way will allow a group of users to easily access all the information relevant to a particular activity or topic. Integrating communications activities into workflow applications would assist in streamlining other daily activities.

Customization and Personalization

Personalization

All of the services and devices above need to be configured in ways that work together effectively. This customization must be flexible and adaptable to various users, communities and groups with which they communicate, and support a variety of devices and applications.

Vision

Communications options have increased to the point where users want to be able to choose what tools they use and how and when they use them. These tools need to be flexible - with options for customizing them to suit an individual's communications style. And the tools need to interoperate with each other in ways that provide a unified communications environment. In the future these tools should be able to inform each other and adapt this environment in real time based in what the individual is doing, where they are, and what they are trying to accomplish.

To support ease in personalization, users should have quick and easy control over this environment. Initially this is most easily provided by self-service web pages that allow users to modify and reconfigure individual systems. A common web interface should be used to configure all a users personal communications options. Devices should also be able to control options that effect the

communications environment, such as activating a “do not disturb” mode for a user.

Users should also be able to easily activate and customize new services via "self service" web pages.

Actions

We should investigate ways to provide common, web based configuration tools that allow users to customize multiple applications. For example, if you are on vacation and will not be reachable you should be able to activate absence messages for voice, email, and IM services from a single interface. This particular function could also be linked to the users calendar and automatically invoked.

While personalization if services should support individuals it should also be possible for groups to customize the services that they use in ways that support their activities and communications styles. A group may want to customize a set of mailing lists, voice call groups, and IM chat rooms based on a set of agreed upon options. Opportunities to integrate user groups, mailing lists, IM and presence rosters should be considered along with the integration of communications activity logs.



Middleware

Middleware components are not seen as specifically related to communications services but common components that support a variety of applications. The desire is to leverage existing components and to only add new ones in this area where they can serve multiple applications.

Vision

Universal identities that span multiple systems will allow for the easy access to both people and services. You should not need to be burdened with having to collect multiple addresses for each person at MIT. Similarly addresses for addressing services should be consistent – for example, the email list address for computing help should also be the same as the SIP address or for an IM chat service.

User have an expectation of secure communications. This includes not only secure media that is encrypted end to end but secure signaling paths for

communications control. Authentication systems should securely identify individuals connecting to the system as well as support federation between organizations. Authorizations for the use of restricted or expensive services should be maintained in one location.

Network management tools should be ubiquitous and not replicated for different types of communications signaling or media.

Actions

We should continue to support a single communications identity for individuals that is used by by all systems for authentication and authorization.

At the point of identifying a SIP soft client that meets other requirements we should consider the options for adding Kerberos authentication to both that client and the SIP infrastructure.

We should improved enhance network management tools where necessary to ensure the quality of real time communications.

Dennis Baron
mailto:dbaron@mit.edu
sip:dbaron@mit.edu