Secure and Mobile Multimedia Convergence

Alex Talevski and Vidyasagar Potdar

Curtin University of Technology
Australia

1. Introduction

In the Information Technology and Telecommunication (IT&T) world, convergence refers to the move towards the use of a single unified interaction medium and media as opposed to the many that we use today. IT&T Convergence aims to enable telecommunications services that are concurrently coupled with enterprise and internet data. The ability to concurrently visualize a concept via sound, images, video, graphs and diagrams while communicating greatly enhances interaction. Communication is more pleasing, meaningful, effective and efficient. This allows for actions to be taken with greater understanding, precision and speed as a response to just-in-time requirements from distributed locations. Therefore, data and telecommunications convergence promises a wide range of possible solutions that will increase productivity and flexibility, and provide new opportunities for modern enterprises. Converged voice and data services have rapidly emerged as a popular alternative to existing telecommunications networks and computer services. Many sources (Phil & Cary, 2009; Stallings, 2004; Deloitte, 2009; Grant, 2005) indicate that converged voice and data networks of various forms are rapidly growing across industry in the last 5 years. However, converged telecommunications and data services have been largely isolated to static environments where fixed Personal Computers (PC) and network connections are used in conjunction with various software tools that simulate pseudo converged sessions. Generally, data presented on the internet and in enterprise applications is not available on telecommunications networks and computer devices and vice-versa. Computer Telephony Integration (CTI), Voice Over Internet Protocol (VoIP) and Interactive Voice Response (IVR) systems form cornerstone technologies behind IT&T convergence. This chapter proposes a secure and mobile multimedia convergence solution.

1.1 Computer Telephony Integration (CTI)

Telephone and computer systems are two technologies that impact many aspects of our daily lives. These technologies drive the world’s economy and are central to the operation of virtually every enterprise. Computer Telephony Integration (CTI) is defined as the integration between computers and telephony systems (Strathmeyer, 1996). CTI technologies bridge the features of computers such as data handling, media processing and graphical user interface with telephone features such as call handling and routing. Currently, CTI is predominantly used to drive software-based Private Automatic Branch eXchange (PABX) systems. However, CTI is heading toward the convergence of both data and voice services over data networks.
1.2 Voice over Internet Protocol (VoIP)
Voice over Internet Protocol (VoIP) (also termed IP Telephony) refers to the transport of voice traffic over data networks. Using VoIP, carrier grade voice communication is digitized and routed in discrete IP packets through a data connection. VoIP is particularly useful when there is limited or financially prohibitive access to alternative telephony networks. Telephone calls can be transmitted with little or no loss in functionality, reliability, or voice quality. VoIP has rapidly emerged as a popular alternative to existing telephony networks (Darlington, 2007; Deloitte, 2009). However, to date, VoIP has been a solution that is mostly used as an alternative medium to carry-out cost-effective long-distance telephone calls.

1.3 Interactive Voice Response Telecom (IVR)
IVR systems provide computer controlled telephone answering and routing functions, as well as facilities for the collection and provision of information. Interactive voice and keypad driven menus allow callers to request information and respond to prompts. Based on developer defined configuration, IVR devices route calls to the appropriate place or system. These systems may use a mixture of human and computer interaction which is provided live, pre recorded by an attendee or digitally synthesized sound to convey information to the caller.

Unfortunately, industry has failed to make the most of CTI, VoIP and IVR technologies. Current use of these technologies has been limited to purely telephony applications like telephone conversations and conferencing. Existing systems rarely provide a flexible and integrated approach where more than telecommunication services are provided. Hence, this chapter proposes a flexible approach that integrates the features of CTI, VoIP, and IVR technologies.

2. Media convergence
Telecommunications and data convergence is required as a consequence of the increased flexibility that businesses demand (Hui & Matthews, 2004). Media convergence aims to enable distributed virtual collaboration environments that would provide all industries and consumers with a new and powerful means for collaboration. From an enterprise perspective, converged telecommunications and data services for its employees, partners and end customers is essential. Converged voice and data services have rapidly emerged as a popular alternative to existing telephony networks. This chapter proposes a foundation for converged voice, video and data solutions. A media convergence solution must have the following key properties;

- **Telecommunications** – Private Automatic Branch eXchanges (PABX) that facilitate call management and routing.
- **Computing** – Graphical User Interfaces (GUI) help visualize interaction through converged telecommunications and computing features such as user interfaces, images, sounds, animations, videos, graphs and tables. Access to enterprise data, applications, services and networks further enhances communication.
- **Convenience** – Traditionally, enterprise applications allow users to access corporate data in a static location using a personal computer. However, accessibility of the proposed feature-rich converged services via a variety of devices in the context of a single converged session offers enterprises great power and flexibility. Such systems must be as convenient as a mobile telephone is for everyday communication.
• **Mobility** - People are no longer desk-bound. Enterprises have to consider the growing population of mobile users that would benefit from the next generation of IT&T services. As more sophisticated wireless devices emerge, the demand for mobile two-way communication will rise dramatically. Flexible, rich access to telecommunications services is crucial in order to achieve optimum performance. New technologies offer innovative features that result in better ways of doing business. Therefore, it is necessary to consider the restrictions imposed by this platform.

• **Functionality** – Current mobile computing devices have powerful hardware, utilize large graphical interfaces and offer network connectivity. With these features, the demand for media and function rich services and multi-way communication on the move will rise dramatically. A novel solution that provides convergence functions on the move is required.

• **Interface** - Desired content must be delivered directly to devices concurrently and interchangeably in a number of formats. Interaction may be performed through various devices and their interfaces (Dual-Tone Multi-Frequency (DTMF), keyboard / mouse, touchscreen, pen etc) and/or through voice driven commands interchangeably. Speech recognition and biometrics (Juang, 1999) aid user interaction.

• **Flexibility** - Due to the diverse nature of this environment, a flexible and adaptive approach is required. However, a problem faced in developing a system such as the one discussed is the complexity of service integration that occurs on the different layers of telecommunications services, telephony networks, computer systems and data networks. Flexible solutions are needed due to the requirement to deploy such solutions in quite diverse roles and environments and unclear, lacking and/or evolving existing enterprise systems. In order to develop a flexible solution that exhibits the required solution properties, a re-configurable service oriented architecture must be employed.

3. **Development approach**

In order to satisfy the key properties outlined earlier, we propose the following features;

**Telecommunications**
- Full Public Branch eXchange (PBX) services
- Telephone calling functions
- Teleconferencing features
- Call forwarding
- Message bank
- Simple Message Service (SMS)
- Call history
- Contacts repository

**Computing**
- Graphical User Interface (GUI)
- Internet
- eMail
- Voice over Internet Protocol (VoIP)
- Enterprise applications
- Instant Messaging (IM) including presence services
Convergence and Hybrid Information Technologies

- Interactive Voice Response Telecom (IVR)
- Global Positioning System (GPS) including location services

Convenience
- The proposed system may be employed to access a variety of telecommunications and computer services via a telephone or computing device.
- Interaction is available via various access devices (TV, PC, PDA, telephone, mobile phone, web and others) both wired and wireless.
- Voice / data transportation mediums (IP, Wi-Fi, Bluetooth, GPS, GPRS, UMTS etc) both wired and wireless connection that is carrier, device and network independent.
- Voice over Internet Protocol (VoIP) including a variety of transportation protocols and encodings (H323, SIP, IAX etc).
- Public Switched Telephone Network (PSTN) and the Plain Old Telephone Service (POTS) services.
- No dependence on specialized hardware devices.

Mobility
- Available and accessible on common telephones, mobile phones and PDAs.
- Conforms to the screen size, processing power, memory, storage and battery life of mobile devices.

Functionality
- Data handling, media processing and graphical user interfaces coupled with telephone features such as call management and routing.
- Full telecommunications and teleconferencing features coupled with whiteboarding, file sharing, instant messaging, meeting and presentation management.
- Customised switchboard intelligent call forwarding and voice mail features are used to manage calls as required.
- Access to enterprise and internet data
- Profiles, contact information, favourites, history
- Voice mail

Interface
- Concurrent streaming media (voice, video, web and data) where interaction can be performed using web, voice, SMS, video, data and instant messaging interfaces.
- Automatic Speech Recognition (ASR)
- Dual Tone Multi Frequency (DTMF)
- Biometrics
- Transparent voice / data switchover

Flexibility
- Proven open technologies with a focus on wide compatibility.
- Re-configurable component-based framework which constitutes the skeletal support that is used as the basis for constructing the solution.
- Simplified software construction, customisation, integration and evolution.
- Complementing components that can operate as a composite or individually.
- Flexible access to telecommunications services and enterprise and internet data.
- Allows prompt awareness and response to enterprise triggers
We aim to promote converged multimedia collaboration in an easy and convenient manner. Therefore, this proposal focuses on an open extensible architecture that uses a mobile thin client approach to allow converged service access via a range of devices and network connections with no specialized hardware or software.

4. Media Convergence Centre (MC²)

Our proposed Media Convergence Centre (MC²) (Figure 1) allows users to participate in a converged multimedia collaboration network using a variety of interaction devices in an easy and convenient manner. A thin client approach is adopted to allow access via a range of devices and connections with no specialized software or hardware. Calls, conferences and data services are provided over wired and wireless telephony and data networks. Interaction can be performed using voice, video, and data streams. The MC² (Figure 1) is composed of the following key components: MC² Communicator, Convergence Centre, Switchboard, Conference, Call, Interact and Services.

![Fig. 1. Media Convergence Centre (MC²)](image)

4.1 MC² Switchboard

A MC² Communicator (Figure 2) focuses on a thin client that is available using a browser and web connection and/or a telephone. As most enterprises have both telecommunications and internet facilities, it is intended to provide collaboration functions just like a telephone provides telecommunications services. The solution aims to facilitate mobile telephony, conferencing, video interaction, instant messaging, SMS and all other MC² services outlined.
below. Furthermore, it is imperative that this client allows users to utilise their call forwarding, contacts, history etc. The MC² Communicator provides access to the MC² Switch board and MC² Conference, MC² Call, MC² Interact and MC² Services via an concurrent and interchangeable Tri-Interface. The solution utilises a number of existing technologies such as JWChat, Ajax, Punjab, Jabber Agi and Asterisk arranged in the following way;

Fig. 2. Media Convergence Centre (MC²)
4.2 MC² Switchboard
The MC² Switchboard (Figure 2) is a web-based tool for managing phone calls. It contains traditional switchboard properties.

4.3 MC² Conference
The MC² Conference (Figure 2) utilizes a thin-client approach. It provides an integrated web conferencing suite. PSTN, GSM and VoIP are integrated in a single session. The solution gives the ability to view and control the status of all participants in a conference. It is possible to interact using voice, video and data. In particular, presentation sharing and flow control as well as file sharing (with version control) are available to enhance the conference experience.

4.4 MC² Call
Full telephony services as per Public Switched Telephone Network (PSTN) and Plain Old Telephone Service (POTS). MC² Call (Figure 2) is available on any browser or telephony device.

4.5 MC² Interact
The MC² Interact (Figure 2) services provide a reconfigurable Interactive Voice Response Telecom (IVR) that enables selected computer controlled telephone answering functions, collection of information and interactive menus for callers to use to input data using the telephone keypad or voice prompts. Based on user defined steps, commands and responses to prompts, calls are routed to a configured place or data. Users can create, change and remove multiple MC² Interact maps. Figure 4 illustrates a high level architecture of the MC² Voice Access to Data (VAD) (Talevski & Chang 2009) solution which provides access to enterprise and internet data. The plugin services used here interact with the user using voice prompts and/or a web interface and access other components, services and repositories as required. Sample vEmail, vFinance, vStocks, vWeather, and vNews plugin services are illustrated.

4.6 MC² Services
To demonstrate the feasibility of such a solution (Figure 2) we developed the following services as IVR plugins. It is possibly to develop virtually any enterprise or internet service as a MC² Interact services. A web interface is provided to personalize and optimize the solution.

- **vEmail Service** - Emails can be accessed at any time and followed up instantly. The vEmail voice plugin hosts a Post Office Protocol (POP) email service where high priority emails can be forwarded for voice access. The vEmail voice plugin reads out each email using a clear voice. The user may interact with the vEmail voice plugin by telephone key tones. MC² Interact allows the automatic browsing of emails without user intervention. It is possible to reply to emails with predefined email templates and forward messages to predefined contacts immediately. Users may manage their email messages by saving, moving and deleting selected items. It is also possible to customize the way that the system performs and to manage emails contacts and template messages via an easy to use web interface.

- **vStocks Service** - Live Australian Stock Exchange (ASX) values can be heard at the user’s convenience. The vStocks service reads out detailed information on each user’s
individually predefined stocks. Stock list navigation is performed using telephone key tones. MC² Interact allows the browsing of stock data without user intervention. The vStocks service is able to announce each stock’s trade date, time, change, previous close, day high, day low, and volume. Users may customize the stock properties they wish to hear to suit their individual preferences.

- **vWeather Service** - Live Bureau of Meteorology (BOM) weather forecasts can be accessed at any time. The vWeather service reads out detailed weather information for a user’s predefined city or town. Weather forecasts are read out for up to one week in advance. Forecast information, days high and days low are given for each day. Users may customize their city to suit their individual preferences and travel arrangements. MC² Interact allows the weekly weather forecast to be read out without user intervention.

- **vNews Service** - Live Rich Site Summary (RSS) News feeds can be heard at a preferred occasion based on a user’s preference. The vNews service reads out each news item as requested by the users telephone key tone interaction or automatically.

An N-Tier distributed platform and component based computing architecture is proposed using the following technologies and tools (Figure 2);

- **Asterisk** - Asterisk (Digium, 2009) is an open source software PBX that can be programmed to create custom applications. Our system uses the java-based Application Gateway Interface (AGI) to trigger custom classes that handle incoming connections.

- **Java** - Java provides a platform independent environment that has wide support. Java (Sun, 2009) was used to interface the tools mentioned and implement the solutions in this system. Enterprise Java Beans are a suitable for the development of distributed and heterogeneous component systems. These technologies were proposed because the Asterisk AGI has wide Java support and many existing libraries and frameworks that interface the soft-switch functions of Asterisk.

- **AT&T TTS** - The AT&T Text-To-Speech (TTS) (AT&T, 2009) engine generates high quality synthesized voice from text. It is integrated through the Asterisk AGI interface. Using the AT&T TTS is it possible to adopt different dialogue files to simulate accents from different nationalities.

- **Speex** - This product (Speex, 2009) provides Automatic-Speech-Recognition (ASR) functions. Asterisk’s extended scripting commands make use of Speex to take voice commands from a user.

- **Hibernate** - Hibernate (Hibernate, 2009) is a high performance object / relational mapping service for Java. It uses interfaces that have defined via mapping documents to convert between the Object Oriented (OO) to the Relational Database Management Systems (RDBMS).

- **MySQL** - In order to persist hibernate objects we us the MySQL RDBMS (MySQL, 2009).


- **AJAX** - Asynchronous JavaScript and XML, is a Web development technique for creating interactive web applications (Ajax, 2009).

- **Punjab** - A HTTP, jabber client interface that allows persistent client connections to a jabber server (Punjab, 2009).

- **IAX** - The Inter-Asterisk Exchange protocol (IAX) (IAX, 2009) was created as an alternative signalling protocol to SIP and H.323. IAX has full signalling and media transfer capabilities that can be used with any type of streaming data (including video). Libiax is a library to take care of the low level network functions. This library was constructed by the makers of Asterisk, and is commonly used by open source IAX clients. The code modifications necessary to support encryption were mostly required within libiax.

- **Cryptlib** - Cryptlib is a powerful, general purpose open-source cryptography package designed to provide security services to applications. Its main purpose is to provide cryptography functions that can be integrated into applications (Cryptlib, 2009).

Clearly, the critical problem in achieving converged IT&T services is being able to combine the many differing technologies and applications that operate computer and telecommunications systems. Furthermore, it is of critical importance that convergence services provide access to enterprise systems.

5. **Service composition**

The ability of systems to adapt or be adapted to disparate enterprise requirements and environmental changes is referred to as their flexibility (Booch, 1994). A flexible system is needed due to the requirement for a system to be deployed in converged enterprise (diversity) and to be flexible to evolving requirements (uncertainty) (Booch, 1994). Versatile systems exhibit generic and function rich properties as a response to growing enterprises’ demand for rapid and frequent development, maintenance, customization and evolution. Convergence needs to facilitate such enterprise growth through integrated telecommunications and computer systems.

Service-based software engineering is a way of raising the level of abstraction for software development so that software can be built by easily reusing previously designed, implemented and refined converged services. Composite architectures that incorporate enterprise services are formed using a Service-Oriented Architecture (SOA) as a standardized way of connecting loosely-coupled systems such as the many that already exist in the computing and telecommunications areas.

Re-configurable service oriented architectures promote simplified software evolution in complex environments. They can be used to provide the glue between enterprise business applications and multi-modal forms of converged business communication to access a variety of telecommunications and data services.

Using a reconfigurable plug and play component-based framework as a basis for the creation and modification of software, it is possible to construct, customize, integrate and evolve convergence solutions in a straightforward way.

5.1 **Service architecture**

The following framework promotes simplified software construction, customisation, integration and evolution of convergence solutions. The framework allows the solution to easily integrate existing enterprise and internet applications and newly implemented components as communication / interaction services. Services may be added and removed dynamically as per business requirements. This allows for prompt awareness and response to enterprise triggers.

At the highest level, MC2 behaves as an IVR entrypoint. As illustrated below (Figure 3), voice plugin discovery, query, identification and invocation are used to situate, define,
describe and utilize available telecommunications and computing services. Once a service plugin has been identified it is assigned access telephone number and appropriate voice/video/data interface. Upon activation, and during execution, each voice plugin governs user interaction and the provision of its converged services.

![Voice Plug-in Diagram](image)

Fig. 3. Converged Service

The plugin architecture defines the following component interface;

- **Discovery** - Plugin discovery is used to find, identify, describe and use available voice plugins. In order to locate plugins, the MC² Interact server broadcasts a request for a voice plugin lookup service. Each voice plugin responds to this request with a lookup proxy.

- **Query** - The MC² Interact host is able to query the voice plugin lookup service for available services.

- **Identification** - The voice plugin lookup service is used to define voice plugin characteristics.

- **Invocation** - When a voice plugin is selected via the IVR the MC² Interact host dynamically binds to the voice plugin and invokes its entrypoint. The voice plugin then takes over interaction control and performs its identified services.

Figure 4 illustrates a high level architecture of the Interact portion of the MC² proposal. It provides access to enterprise and internet data using an Interactive Voice Response (IVR) system and concurrent web interface. The plugin services used here interact with the user using voice prompts and/or a web interface and access other components, services and repositories as required. Sample vEmail, vFinance, vStocks, vWeather, and vNews plugin services are illustrated.

### 6. Security

Corporate customers are generally more security conscious. They require that potential new technologies are proven not to be a security risk. Most current converged and VoIP-based offerings do not offer a practical security solution. However, an important aspect behind the corporate success of the telecommunications and data convergence is security. As these technologies become more heavily integrated into the workplace, so too do the
Fig. 4. MC^2 Interact

opportunities for hackers. Converged information is generally routed unsecured through data packets on a public network. There is software that can capture, reconstruct and/or modify these sensitive interactions which opens numerous security concerns as follows (Bilby, 2009):

- Eavesdropping and recording phone calls
- Tracking calls
- Stealing confidential information
- Modifying phone calls
- Making free phone calls
- Pranks / Practical jokes
- Board room bugging
- Sending spam (voice or email)

There are currently very few practical security standards available to secure converged telecommunications and computing services on mobile devices.. Furthermore, many enterprises that have adopted converged technology have not been able to effectively secure these solutions as a result of multi-vendor incompatibilities (Gohring, 2009). To alleviate this, it is necessary to add some form of protection, such as encryption, at the transport or network layer. By incorporating security at each level of the network, it makes successful attacks much more difficult. Simply breaking one type of security will not expose the entire
network; it would require multiple levels of protection to be compromised. However incorporating multilevel security would incur additional cost, which should be considered. To allow multi-vendor solutions to interoperate it is essential that such solutions are integrated into a mobile convergence standard. There have been many attempts to provide secure services for the major convergence protocols (Abad, 2003; Arkko & Carrara, 2004; IAX, 2009). Unfortunately, these systems typically suffer from the following problems:

- Complicated to deploy and maintain
- Rely on proprietary and/or incompatible solutions
- Require an existing Public Key Infrastructure (PKI) and/or other resources
- Experience Significant routing problems when passing through NAT

6.1 Encryption algorithms

In order to provide secure transmission of data, it is necessary to offer confidentiality and authentication. In other words, data must be valid and should not be available nor disclosed to unauthorized parties. In order to support different codecs, the encryption algorithm must be able to support variable length data payloads where the amount of data per frame is likely to be short but send at a high frequency (approximately 30-100 bytes 50 times per second). As the data payload is relatively small, it would be advantageous to use an encryption method that will not increase the size of the data to be sent. Any small increases in size will add significant overhead to the transmission.

The Inter-Asterisk Exchange protocol (IAX) was created as an alternative signalling protocol to SIP and H.323. It provides full converged media transfer capabilities. The block diagrams below give a basic description of the structure of the IAX software layers and an added security layer (Figure 5). After the converged data has been encoded, it is intercepted and encrypted before being sent across the network. At the receiver’s side, the data is decrypted, and passed back through the normal IAX processing stack.

![Fig. 5. Modified IAX Architecture Diagram](image)

7. Conclusion

IT&T convergence refers to the move towards the use of a single united interaction medium and media as opposed to the many that we use today. Such convergence aims to enable telecommunications services that are concurrently coupled with enterprise and internet data. This newfound visibility greatly enhances interaction through sound, images, video,
graphs and diagrams. Communication is more pleasing, meaningful, effective and efficient. Therefore, enterprises can take actions as a response to market drivers much more quickly and with greater precision. Data and telecommunications convergence promises a wide range of possible solutions that will increase productivity and flexibility, and provide new opportunities for modern enterprises. This chapter proposed a secure and mobile multimedia convergence solution.

8. References
Deloitte, “Getting off the Ground: Why the move to VoIP is a decision for all CXOs”, On-line at: http://www.deloitte.com/dtt/budget/0,1004,cid%253D67263,00.html (2009).
Starting a journey on the new path of converging information technologies is the aim of the present book. Extended on 27 chapters, the book provides the reader with some leading-edge research results regarding algorithms and information models, software frameworks, multimedia, information security, communication networks, and applications. Information technologies are only at the dawn of a massive transformation and adaptation to the complex demands of the new upcoming information society. It is not possible to achieve a thorough view of the field in one book. Nonetheless, the editor hopes that the book can at least offer the first step into the convergence domain of information technologies, and the reader will find it instructive and stimulating.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: