Converged Voice Access to Data (CVAD)

Alex Talevski, Elizabeth Chang and Tharam Dillon

DEBI Institute, Curtin University, GPO Box U1987, Perth, WA, 6845, Australia
e-mail : (alex.talevski, elizabeth.chang, tharam.dillon)@cbs.curtin.edu.au

Abstract

The rapid growth of the Internet and widespread use of mobile computing and telephony has resulted in a broad range of ways to communicate and access information. Data and telecommunications convergence promises a wide range of possible solutions that will increase productivity, reduce costs, and provide new opportunities and revenues for enterprises. However, such converged telecommunications and data services have been largely isolated to static environments where fixed Personal Computers (PC) and an Internet connection are used in conjunction with various software tools to simulate a pseudo converged session. Generally, data presented on the internet and in enterprise applications is not available on voice networks and vice-versa. Due to the versatile nature of today's enterprise, a feature-rich, flexible, adaptive and widely accessible convergence solution is required.
This paper presents an Interactive Voice Response (IVR) driven approach that uses Converged Voice Access to Data (CVAD) services. The CVAD solution offers enhanced service effectiveness, flexibility and convenience for professionals on the move.

1. Introduction

The advent of personal computers and the web have mostly provided distributed mechanisms for information sharing while telephones serve almost exclusively as communication devices [1, 2]. Currently, enterprises utilize a heterogeneous mixture of:

- **Data and voice networks (wired and wireless)** - Data and voice networks are used to perform daily tasks. Typically, Public Switched Telephone Network (PSTN), Global System for Mobile (GSM) and Internet Protocol (IP) networks provide direct access to telecommunications and data services.

- **Access devices -** Wireless and wired telephones and computing devices are used to provide links to data and voice networks and services.
- **Data repositories and applications** – Data repositories and applications store and process enterprise information and provide optimized interfaces and views of this data.

In the IT&T world, 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 [3]. Such a solution would enable telecommunications services that are concurrently coupled with enterprise and internet data. The ability to visualize a concept via images, graphs, tables and procedures while communicating over the telephone greatly enhances interaction. Interaction is more pleasing, meaningful, effective and efficient. Feature-rich services allow for actions to be taken on the move with greater precision and faster response to market drivers. Therefore, data and telecommunications convergence promises a wide range of possible solutions that will increase productivity and flexibility, reduce costs, and provide new opportunities and revenues for enterprises.

Converged voice and data services have rapidly emerged as a popular alternative to existing telecommunications networks and computer services. Many sources [4 - 7] indicate that converged voice and data networks are growing from only 100,000 US households in 2004 to more than 12 million by 2009. 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 voice networks and devices and vice-versa.

From an enterprise perspective, converged voice and data services and interaction mediums to its employees, partners and end customers is essential.

Due to the diverse nature of this environment, a Converged Voice Access to Data (CVAD) solution that employs the features outlined above is required. This paper proposes an environment that has the following features;

- **Convergence** - Converged voice, video and data services and interaction mediums are required as a consequence of the increased flexibility that businesses demand [8]. Such a collaboration environment is more pleasing, meaningful, effective and efficient.

- **Accessibility** – The demand for media and function rich services and multi-way communication on the move is rising. The availability of converged telecommunications and, enterprise and internet services via a variety of mobile devices is essential. Such devices (TV, PC, PDA, Mobile Phone and others), voice/data transportation mediums (IP, Wi-Fi, Bluetooth, GPRS, UMTS etc) and a variety of protocols and encodings (H323, SIP, IAX etc) must be considered.

- **Feature Rich Services** – The combination of major telecommunication functions such Interactive Voice Response (IVR) with services that access enterprise and internet data such as databases, applications and web services via multiple interfaces. Access to enterprise data repositories, applications (Databases, ERP, CRM and others) and Internet data services (Email, News, Weather, Stocks and others) is essential.

- **Multi-Interface -** Being able to visualize the context of a conversation greatly enhances interaction. Voice, video and data interfaces project service data and system/user interaction. This paper proposes system interaction that is interchangeably performed through the device interface (DTMF, keyboard / mouse, etc) and/or through voice driven commands.

- **Flexible and Adaptive Environment** - A problem faced in developing a feature-rich, flexible and widely accessible solution is the service integration and reconfiguration issues occurring on the disparate layers of telecommunications and computer services. Re-configurable software architectures promote simplified software evolution in complex environments.

1.1. Solution Background

1.1.1 Computer Telephony Integration (CTI).
Computer Telephony Integration (CTI) is defined as the integration between computers and telephony systems [1]. CTI is a bridge that connects features of computers such as data handling, media processing and graphical user interface with telephone features such as call management and routing. Currently CTI is predominantly used to drive software-based Private Automatic Branch eXchanges (PABX).

1.1.2 Voice over Internet Protocol (VoIP).
Voice over Internet Protocol (VoIP) also termed IP Telephony is a solution that is commonly used by people as an alternative medium to carry-out telephone calls [8]. Generally, VoIP refers to the transport of voice traffic over a data network. VoIP hardware and software acts as an Internet transmission medium for telephone calls and other converged data [9]. Telephone calls can be transmitted with little or no loss in functionality, reliability, or voice quality.

1.1.3 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 menus allow callers to input data using the telephone keypad and/or simple voice prompts. IVR devices route calls to the appropriate place or data, based on user defined steps, commands and responses to prompts.

1.1.4 Service Oriented Architecture (SOA).
Service-Oriented Architecture (SOA) [10] embraces an architectural approach to developing composite applications from prefabricated, loosely coupled and collaborating agents. SOAs encourage larger applications to be divided into smaller discrete modules. These fine-grained modules are used to produce course grained services that can be easily integrated by others. SOA defines how service interfaces are located, executed, managed, monitored and secured. Using SOAs dynamic composition, integration and tailoring of converged media services is possible [11].

Industry has failed to make the most of CTI, VoIP, IVR and SOA technologies. Concurrent use of these technologies in the Information Technology and Telecommunications (IT&T) world has been limited to purely telephony applications like telephone conversations and conferencing. Existing systems rarely provide an integrated approach where more than communication services are provided. A flexible approach that integrates the best features of CTI, VoIP, IVR and SOA technologies is required.

Section 2 outlines Converged Voice Access to Data (CVAD). Sections 3 and 4 detail the adopted reconfigurable plugin framework and Meta model. Section 5 details the implementation, use of third party tools and existing plugins. Section 6 concludes the paper.

2. Converged Voice Access to Data (CVAD)

The CVAD system eliminates traditional barriers by offering a new and novel solution that allows flexible access to a range of enterprise and internet services on the move. It provides the glue between enterprise business applications and multi-modal forms of business communication. The system is employed to access a variety of telecommunications and data services with or without the use of a personal computer or an internet connection. Figure 2 illustrates a conceptual CVAD system architecture with sample services (vEmail, vFinance, vStocks, vWeather, and vNews). The services are business modules that interact with the user using voice prompts and access other components, services and repositories as required.

3. Reconfigurable plugin framework

A problem faced in developing such a CVAD solution such as the one discussed earlier is the service integration complexity [12, 13] and configuration occurring on the telecommunications, computer systems and networks layers. A reconfigurable plugin framework must be utilized to customize, integrate and subsequently evolve convergence solutions. An Interactive Voice Response (IVR) system is programmed using a reconfigurable plug and play software architecture [12] to dynamically compose, integrate and tailor voice access to data services. This RVAD architecture promotes simplified software

construction, customization, integration and evolution. It constitutes the skeletal support that is used as the basis for constructing composite component architectures from interconnected components. The architecture makes it possible to easily integrate existing enterprise and internet applications and newly implemented components as voice plugins. Customized voice plugins are constructed by mixing and matching voice components, interaction engines and data source components. Once constructed, voice plugins expose an IVR entrypoint. Desired content is then delivered direct to devices in an audible format within the context of a single converged telephone conversation. System interaction is performed using voice commands or touch-tones. Once a voice plugin has been identified, user commands invoke the voice plugin's data services. During execution, each voice plugin governs user interaction and the provision of its services. As illustrated below (Figure 1), voice plugin discovery, query, identification and invocation are used to situate, define, describe and utilize available services.
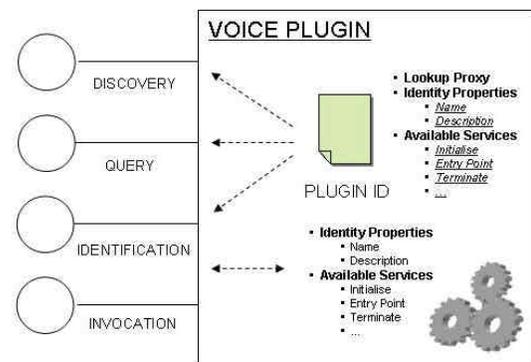

Figure 1. Voice plugin
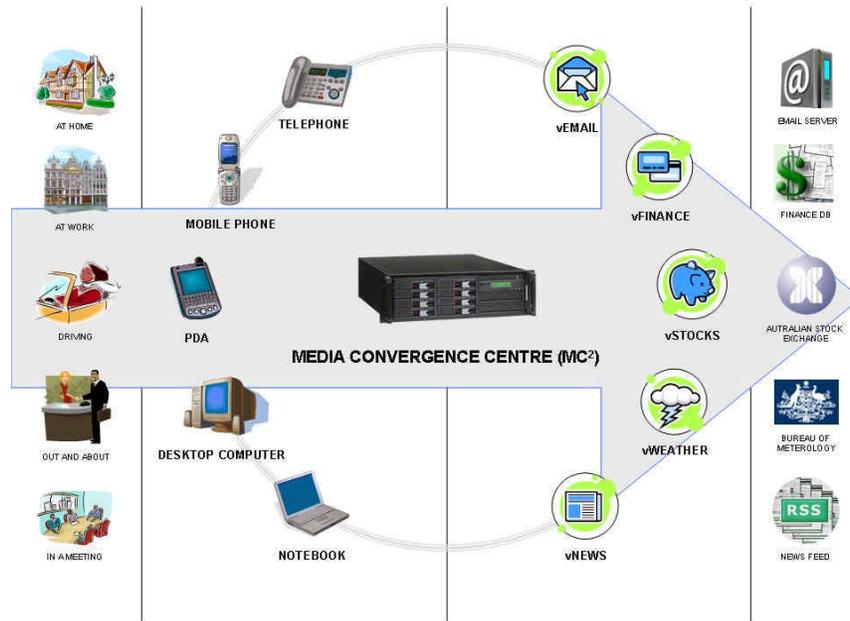
Figure 2. Converged Voice Access to Data (CVAD)

3.1 Application solutions

The CVAD application (Figure 3) is constructed from a number of lower level voice plugins that are fully composed business modules. The diagram below illustrates the high level CVAD solution with a set of sample business modules.
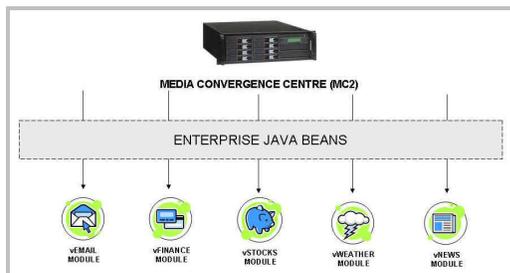
Figure 3. Application Solution

3.2 Business Modules

Business modules (Figure 4) represent voice plugins that typically perform a single voice access to data service. Such business modules utilise groups of component compositions. The high-level diagram below illustrates the vEmail business module. This business module is constructed from multiple lower-level generic components which are reused in other business module compositions.
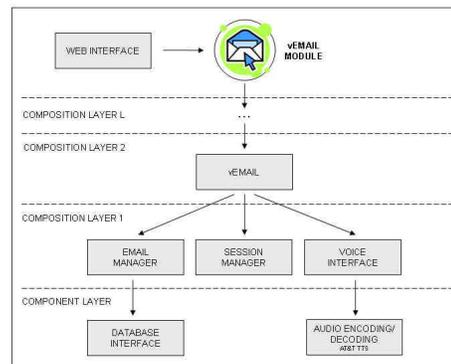
Figure 4. Business Module

4. Meta Model

A software component may expose a number of interfaces. Each interface either provides or requires services in the form of attributes and operations. An attribute is a named property value that defines the characteristics and state of a component. An operation is the implementation of a specific service that represents the dynamic behavior of a component. Operations are specified using their input and output parameters and pre and post conditions. The component Meta data model illustrated in Figure 5 represents an explicit Meta specification of a single software component and its service interface.
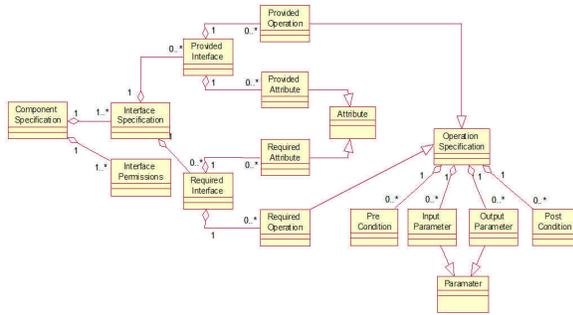
Figure 5. Conceptual Meta Model of a Component

A composition is typically formed from many interconnected components that are constructed in a layered and hierarchical manner. Connectors provide a level of indirection that reduces dependencies among components. Interconnected compositions are coupled using operation and attribute connectors and adaptors.

5. Implementation and third party tools

The software platform is an N-Tier distributed computing architecture that employs Enterprise Java Beans (EJB), Java 2 Enterprise Edition (J2EE), and XML technologies. The CVAD server IVR utilises an interaction Finite State Machine (FSM) to govern voice access to plugin services. The System can be broken into two major subsystems.

- The CVAD server and plugins.
- Web based configuration.

5.1 CVAD server and plugins

- Login and session management is provided by an authentication component. This component provides authentication and session validation functions.
- A voice interface component is responsible for all user interactions.
- Audio encoding/decoding is performed by a voice data component.
- Voice XML (VXML) is encoded/decoded using a voice document component. This component is responsible for the creation, parsing and programmatic manipulation of voice documents
- A database interface component stores all persistent input and output data.

5.2 Web based configuration

- Login and session management provides authentication and session validation functions for web users.
- A user-friendly web user interface is responsible for all online interaction.
- A database interface component performs all relevant input and output persistence.

The solution utilizes a number of third party tools. The most significant tools are outlined below.

- **Asterisk** – Asterisk [14] is an open source software PBX that can be programmed to create custom applications. Our system uses the Asterisk Application Gateway Interface (AGI) to trigger custom Java classes that handle incoming connections.
- **AT&T TTS -** The AT&T Text-To-Speech (TTS) [15] 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 [16] provides Automatic-Speech-Recognition (ASR) functions. Asterisk's extended scripting commands make use of Speex to take voice commands from a user.
- **Hibernate** – Hibernate [17] 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 [18].
- **Java -** Java provides a platform independent environment that has wide support. Java [19] is used to interface the tools mentioned and implement the solutions in this system.

5.3 Existing plugins

vMail Service - Emails can be accessed at any time and followed up instantly. The vMail plugin hosts a Post Office Protocol (POP) email service where high priority emails can be forwarded for voice access. The user may interact with the vMail plugin by telephone key tones. However, CVAD 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.

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. CVAD 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.

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. CVAD 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.

6. Conclusion

Widespread use of mobile computing devices, telephones along with Internet services has resulted in a broad range of ways to access information and to communicate. However, true access voice and data convergence has been isolated to certain environments. Converged services that are accessible anywhere and anyhow promise increased productivity, reduce costs, and new opportunities and revenues for enterprises. A versatile approach is required to produce a solution to this problem. This paper presents an Interactive Voice Response (IVR) driven solution that uses a reconfigurable plug and play software framework and platform to dynamically compose, integrate and tailor voice access to data services. This combination reduces development and integration complexity, effort and cost while enhancing service effectiveness, flexibility and convenience for professionals on the move.

7. References

[1] C. R. Strathmeyer, "An Introduction to Computer Telephony", IEEE Communications Magazine, 35(5), May 1996, pp. 106–11.
[2] B. Benner, "Computer Telephony Integration (CTI) Industry", On-line at: http://faculty.ed.umuc.edu/~meinkej/inss690/benner/CTIpaper.htm (2006).
[3] R. Darlington, "What is multimedia convergence and why is it so important", On-line at: http://www.rogerdarlington.co.uk/Multimediaconvergence.html(2007)
[4] S. Phil, F. Cary, *You Don't Know Jack About VoIP*, Queue, 2004, 2(6), p. 30-38.
[5] W. Stallings, *Data and Computer Communications* (Seventh Ed.), Pearson Educational International, 2004.
[6] Deloitte, "Getting off the Ground: Why the move to VoIP is a decision for all CXOs", On-line at: http://www.deloitte.com/dtt/research/0,1015,sid%3D2245&cid%3D64027,00.html (2004).
[7] M. Grant, "Voice Quality Monitoring for VoIP Networks", Calyptech Pty. Ltd., Melbourne, 2005.
[8] M. C. Hui, H.S. Matthews, "Comparative analysis of traditional telephone and Voice-over-Internet Protocol (VoIP) systems", In Proceedings of the IEEE International Symposium on Electronics and the Environment, 2004
[9] B. Goode, "Voice over Internet protocol (VoIP)", Proceedings of the IEEE, Sept. 2002
[10] "Service-oriented architecture (SOA)", On-line at: http://www.service-architecture.com/web-services/articles/service-oriented_architecture_soa_definition.html (2003).
[11] A. Talevski, E. Chang, T.S. Dillon, "Re-Configurable Web Services for Extended Logistics Enterprise", IEEE Transaction on Industrial Informatics , vol. 1, no. 2, May 2005, pp. 74-84.
[12] G. Booch, "Software Architecture and UML", On-line at: http://www.rational.com/products/rose/whitepapers.jsp (2003).
[13] I. Sommerville, G. Dewsbury, K. Clarke, M. Rouncefield, "Dependability and Trust in Organisational and Domestic Computer Systems", In Trust in Technology: A Socio-technical Perspective, Kluwer 2004.
[14] "Asterisk Home", On-line at: http://www.asterisk.org/ (2006)
[15] "AT&T TTS", On-line at: http://www.research.att.com/viewProject.cfm?prjID=315 (2006)
[16] "Speex Home", On-line at: http://www.speex.org/ (2006)
[17] "Hibernate Home", On-line at: http://www.hibernate.org/ (2006)
[18] "MySQL Home", On-line at: http://www.mysql.com/ (2006)
[19] "Java Sun Home", On-line at: http://java.sun.com/ (2006)