VoiceXML and VoIP

Rob Marchand
Genesys Telecommunications Laboratories Inc.
August 7th, 2006
Agenda

• The Rise of Open Systems
• State of the Nation
  – Dialog Related Standards
  – VoIP Standards
• A Peek at the Future
• Q&A
The Rise of Open Voice Systems

Steady decline in traditional IVR sales accompanied by a steady increase in sales of VoiceXML platforms with speech.
Cost of Staying with the Old Paradigm

- Companies are faced with aging systems that require a significant investment to replace/upgrade
- Increasing difficulty and expense in finding specialized resources to handle:
  - Proprietary development
  - Maintenance and troubleshooting
  - Installation
- Difficulties in integrating to customer backend and CRM systems
- Lack of a tightly integrated contact management solution (self/assisted service) creating inefficiencies in handling customer interactions
The Advantage of the New Paradigm

• Open systems allow for choice in selection of hardware/software
• Systems software upgradeable, reducing overall upgrade costs
• Standard, off the shelf hardware supportable by a wider pool of IT resources and VAR/SI vendors
• Open development environment provides access to larger developer base, and leverages expertise of in-house web developers
• Overall infrastructure can be converged and managed through the company’s IT department
Compelling Events Driving the New Paradigm

VoiceXML

• Evolution of a standards-based programming language for voice
• Web-based development bringing the web paradigm to voice applications
• Drive to openness of voice systems, including use of standard off-the-shelf hardware and standard protocols

Voice Over IP

• Standardization of packet voice
• Convergence of voice and data
• Openness of systems, and greater interoperability between vendor technologies
• Features above and beyond anything available in TDM telephony
Evolution of Standards-Based Voice Processing

- **VoiceXML 1.0** created by AT&T, Lucent, Motorola and IBM (March 2000)
- **VoiceXML 2.0**, **SRGS 1.0**, released by W3C (March 2004)
- **SSML 1.0** released by W3C
- **SISR 1.0 WD** released by W3C
- **PLS 1.0 WD** released by W3C
- **VoiceXML 2.1 CR**
  - **CCXML 1.0 LCWD** released by W3C
- **SCXML 1.0 WD**
- **VXML 3.0 / CCXML 2.0**
  - **FWD** (~2007)
- **SSML 1.0** released by W3C (Sept 2004)
- **PLS 1.0 WD** released by W3C (Feb 2005)
- **SISR 1.0 WD** released by W3C (Nov 2004)
- **VoiceXML 2.1 CR**
  - **SCXML 1.0 WD** released by W3C
- **SPCM**
  - **VXML 3.0**
  - **CCXML 2.0**
  - **FWD** (~2007)
- **Speech Synthesis Markup Language**
- **Semantic Interpretation for Speech Recognition**
- **Pronunciation Lexicon**
- **Call Control XML**
- **State Chart XML**

- **February 2005**
- **PLS 1.0 WD** released by W3C
- **November 2004**
- **SISR 1.0 WD** released by W3C
- **September 2004**
- **SSML 1.0** released by W3C
Standards Progress Overview

• Dialog Standards are well understood
  – VoiceXML 2.0 recommendation has been stable for 2 years
  – Supported by many vendors
  – 15 companies certified, 3 have multiple products!
  – Widely deployed in production
• Continued Evolution
  – VoiceXML 2.1 - Candidate Recommendation since June 13th, 2005
  – Adds useful features orthogonal to VoiceXML 2.0
• But!
  – W3C doesn’t address the plumbing for VoIP
  – A job for the IETF!
The Advent of Voice Over IP

- Leverages converged network infrastructure using a common protocol (TCP/IP) for all media and data types
- Standard protocols (ex. SIP) create cost savings and efficiencies:
  - Disaggregation of components based on function
  - Large market of best of breed vendor solutions at competitive prices
  - Vendor interoperability ensures flexibility in vendor choice:
    - Multiple vendors can be selected for a specific implementation
    - Maximize feature capabilities while minimizing cost
- A wide range of new features can be enabled beyond what is possible using TDM technologies
SIP & Voice Platforms

- SIP has established itself as the preferred network interface to voice platform products
  - PSTN, H.323, other access methods being handled via “gateways”
  - Consistent model for Service Provider and Enterprise
- RFC 3261 governs session establishment and termination, but says little about semantics of a session
  - Telephony usage of SIP well understood (From, To, etc)
  - RFCs and Internet-Drafts exist covering functionality beyond basic session establishment (e.g. REFER – RFC 3515)
  - Voice/Multimedia Platform use cases not widely covered!
SIP Standards for Voice Platforms

• NETANN (RFC 4240) defines basic media services, and how to access those services via SIP
  – De-facto standard for some time, but RFC only as of December 2005
  – Originated from traditional/telco media server space, but applicable to VoiceXML voice/multimedia platforms

• NETANN defines three services:
  – Simple announcement service
  – “Prompt and collect” service – for launching VoiceXML dialogs
  – Basic conferencing service
SIP Standards for Voice Platforms

- Internet-Draft draft-burke-vxml builds on RFC 4240
  - Specifically focuses on expanding NETANN dialog service
  - Two-way information passing to/from VoiceXML platforms
  - Also specifies semantics of variables, transfers, outbound calls
- Provides a generic service for any SIP-based architecture
  - Gateway can invoke service by provisioning appropriate URI
  - SIP AS can use semantics for finer control of VoiceXML dialog
  - CCXML control over dialog provided by this interface
  - “Web service” for the telecommunications network!
- Covers elements beyond setup and teardown
  - VoiceXML platform initiated outbound calls
  - Call transfers via REFER
  - Media behavior (offer/answer, early media) also specified
- Don’t forget MRCPv2, MSML/MSCML/MSCP, others!
A Peek at the Future
Enterprise IVR Architecture

- VoIP-based architectures for dedicated/standalone enterprise IVR becoming increasingly commonplace
- Early IP PBX/ACD-based architectures are much more rigid by comparison
- Newer versions of IP PBX software have support for SIP, but remain fairly closed in terms of overall system architecture
- Voice application delivery is increasingly done on commodity hardware with standards like VoiceXML, CCXML
- The same infrastructure will support multimodal/multimedia specifications
SIP & CTI – Current Model

• CTI address key limitations of PSTN data passing
  – Also provides single view of a call across entire lifetime of that call
  – Can’t we pass all that data with the call in SIP?

• Enterprise IP PBX infrastructure is still highly limited
  – Supports SIP…
  – … but limited information passing – generally no more than what was already available with PSTN/H.323 protocols

• IVR Platforms thus still require proprietary CTI interfaces
  – Even ANI/DNIS otherwise sometimes unavailable
  – Transfers & agent routing with screen pop otherwise not possible
  – SIP-based IVR may need to transfer to PSTN-based contact center!
SIP & CTI – Future Model

• “Ideal” future environment is 100% SIP-based with full information passing between all elements that touch a call
  – End to hunger globally, worldwide peace also part of this vision

• However, CTI integrated transparently via SIP allows effective hybrid deployments that leverage SIP
  – Standards such as draft-burke-vxml can be used to deliver information to IVR platforms without CTI integration…
  – CTI infrastructure can manage routing & information passing to and within legacy environments
  – Vision behind existing Genesys SIP Server product
Summary

• SIP is key to voice as a service in any architecture
  – “Voice” really means multimedia interaction in this context
• Practical benefits to be reaped today, despite any issues
  – Numerous customers already benefiting significantly
• Evolution of CTI will integrate elements via SIP
  – CTI will use SIP services; not SIP services invoking CTI
• Multimedia, multimodal, presence, others built on SIP
  – Despite common use as telephony replacement, many unique new kinds of applications are enabled in SIP-based environments
Empowering Your customers and employees with speech technologies

SpeechTEKI 2006
The Voice Solutions Showcase

Empower

Thank You