VoiceXML and VoIP

Architectural Elements of Next-Generation Telephone Services

RJ Auburn

***Chief Network Architect,* Voxeo Corporation**
***Editor, CCXML Version 1.0,* W3C**

Ken Rehor

***Software Architect,* Nuance Communications**
***Chair, Conformance Committee,* VoiceXML Forum**
***Chair, Voice Browser Interoperation subgroup*, W3C**
***Editor, VoiceXML Version 2.0,* W3C**

VoiceXML Planet Conference   San Jose, California USA   2002 March 21

Overview

- **VoIP Overview**
  - Connection Protocols
  - Audio Protocols
- **Voice Application Deployment Architecture**
  - PSTN
  - VoIP (SIP)
- **VoIP advantages**
  - Flexible Network Topology
  - Complex call routing
- **CCXML**
  - 3^{rd} Party Call Control
  - VoiceXML as a Network Resource
- **Voice Browser Interoperation**

VoIP Overview

- **Connection Protocols**
  - SIP, H.323
- **Media Protocols**
  - RTP, RTCP, RTSP

VoIP Connection Protocols

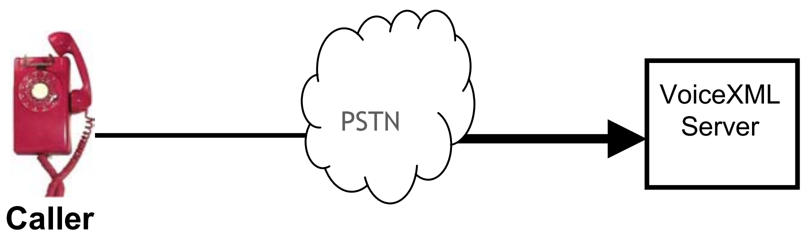
- **Session Initiation Protocol**
  - Lightweight, extensible
  - Based on HTTP and SMTP
  - Developed in IETF
  - Latest draft spec: RFC2543
    http://search.ietf.org/internet-drafts/draft-ietf-sip-rfc2543bis-09.txt
  - See http://www.jdrosen.net/
- **H.323**
  - Originally designed for audiovisual conferencing
  - Popular with VoIP audio-only media connections
  - Developed in ITU

VoIP Media Transport Protocols

- **RTP – Real Time Protocol**
  - Developed in IETF
  - RFC 1889
  - Latest draft http://www.ietf.org/internet-drafts/draft-ietf-avt-rtp-new-11.txt
- **RTCP – Real Time Control Protocol**
  - Developed in IETF
  - Compliment to RTP
- **RTSP – Real Time Streaming Protocol**
  - Developed in IETF
  - RFC 2326

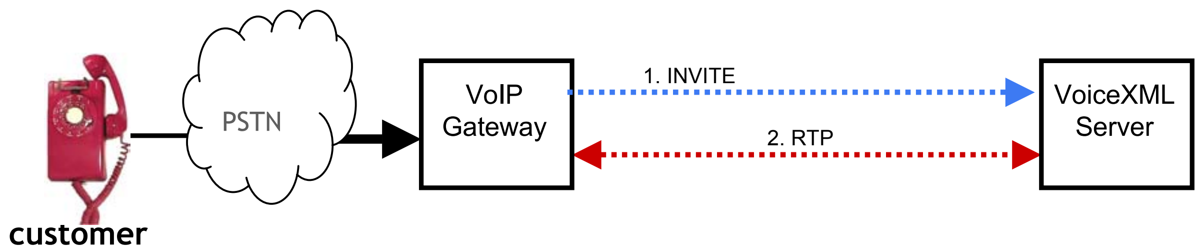
Inbound call using PSTN connections

- **VWS handles call routing/setup/answer**



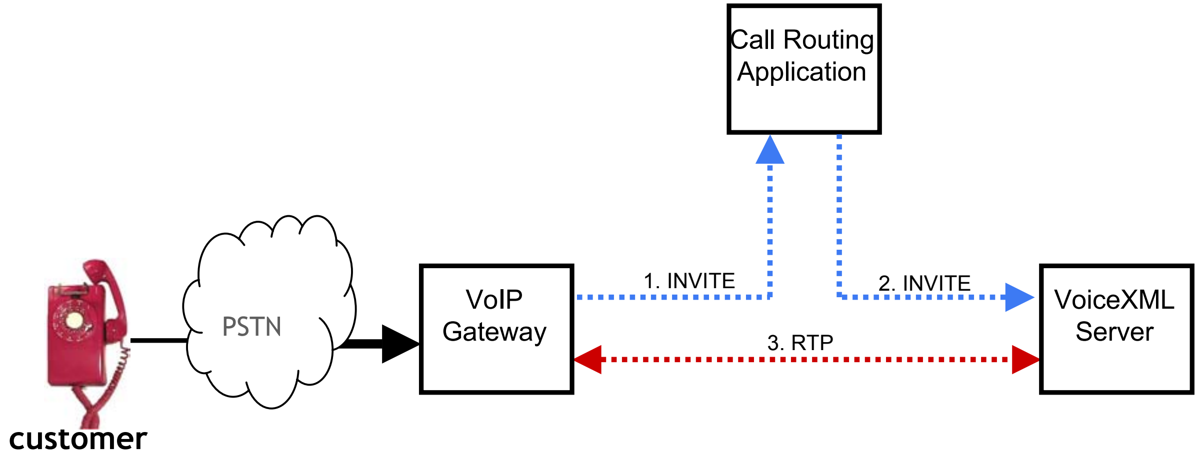
Inbound call using VoIP (SIP and RTP)

- **VWS handles call routing/setup/answer**



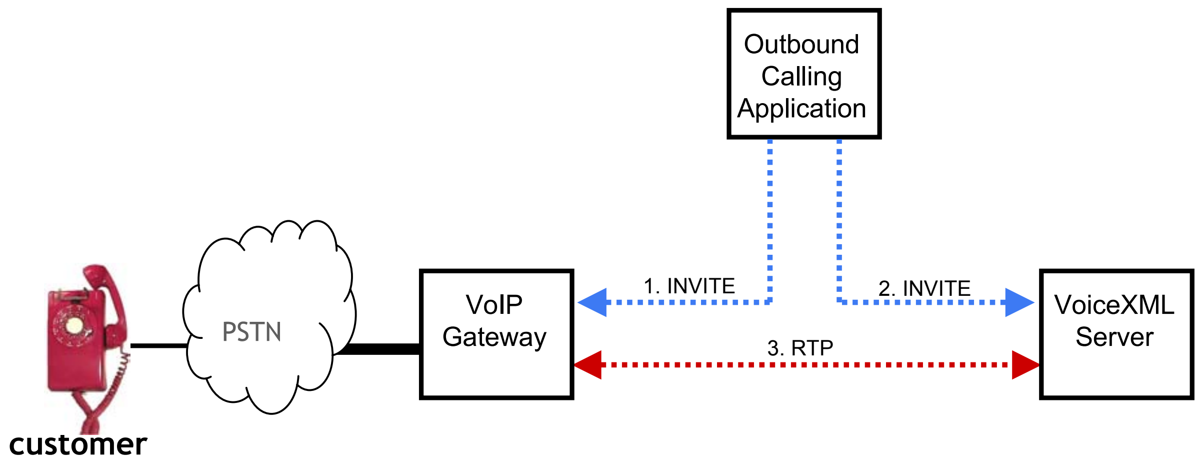
Inbound call using VoIP (SIP and RTP)

- 3rd party application handles call routing/setup/answer



Outbound call using VoIP (SIP and RTP)

- 3^{rd} party call control application



Call Control XML (CCXML)

Call Control For VoiceXML Applications.

CCXML Overview

- **What is CCXML?**
  - Is an XML based language that can control the setup, monitoring, and tear down of phone calls
  - Provides a set of high level tags that support most call control features.
  - Allows for easy integration into back end web applications very similar to VoiceXML's model.
  - Uses the finite state machine model. You write event handlers to move from one state to the next using markup tags.
  - CCXML also provides commands to run a "Dialog" on a call leg (Eg. VoiceXML)

CCXML Overview

- **Why is CCXML Needed?**
  - VoiceXML was designed to be a dialog markup language not a Call Control markup language.
  - VoiceXML does have the <transfer> element but it is limited in it's feature set.
  - Call Control requires advanced asynchronous event handling.
  - Adding Call Control to VoiceXML can be done but it complicates the language greatly.
  - VoiceXML has standalone uses outside of a telephone.

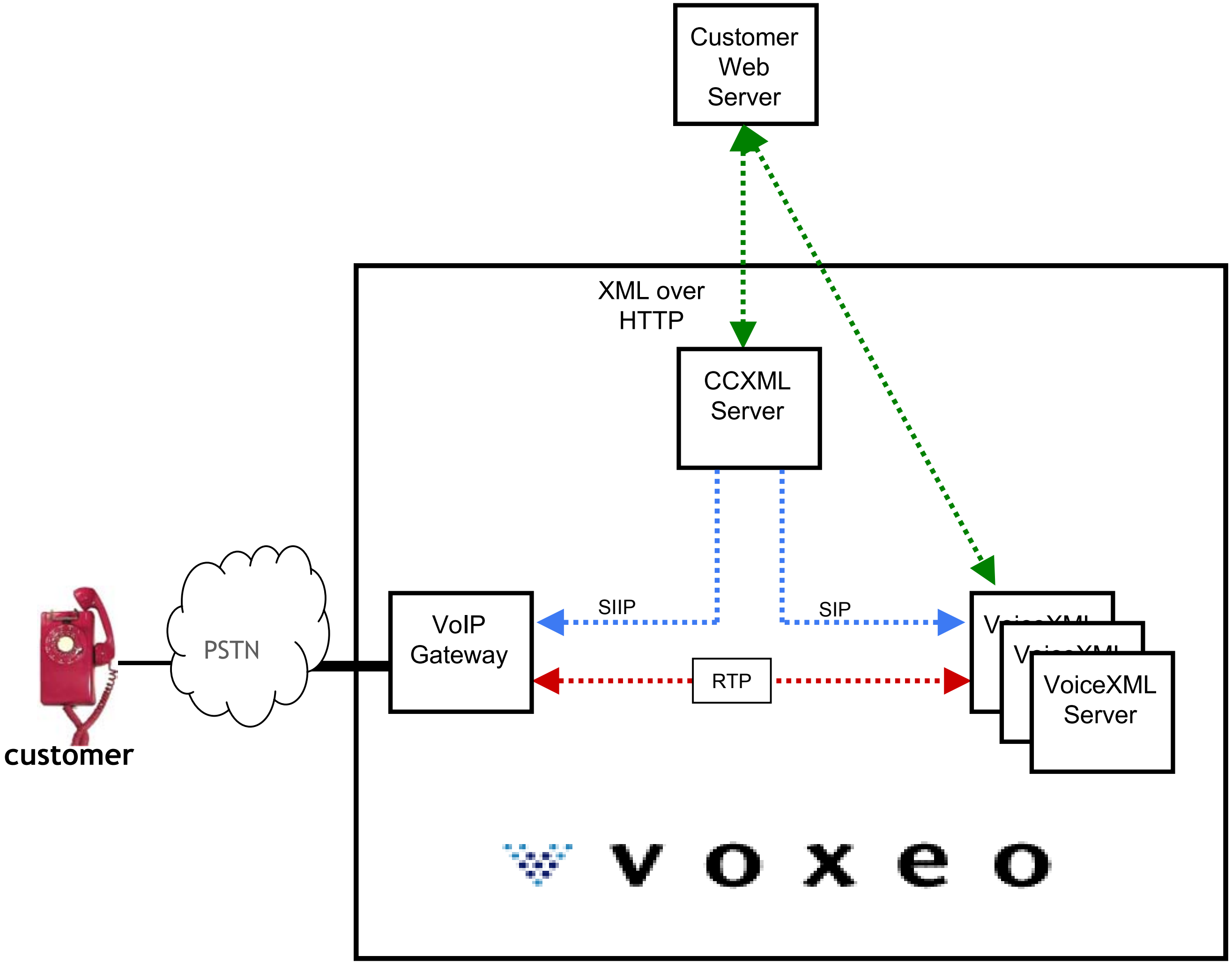
CCXML Overview

- **Why VoIP?**
  - A CCXML platform can integrate with VoiceXML platforms very easily using VoIP based technologies
  - Allows vender interoperation between different VoiceXML platforms and CCXML platforms.
  - A VoIP based architecture allows you to mix and match the best of breed technologies to create better solutions (Eg. VoiceXML, Conference and VoIP Venders)

CCXML Overview

- **CCXML Supports Multiple Dialog Systems**
  - VoiceXML (CCXML was designed around integrating with this)
  - SALT (Integrated via the <smex> tag)
  - CallXML
  - Traditional IVR Systems
- **CCXML Integrates with Next Gen Platforms and Networks**
  - VoIP: CCXML works great with VoIP based platforms. CCXML can sit in a “Soft switch” type role connected to multiple Voice Browsers.
  - TDM: CCXML Could also be integrated in a more traditional card based platform. The CCXML interpreter would switch calls using the card based API’s.

CCXML System Architecture - VoIP



CCXML and Voice Dialogs (VoiceXML)

- **Dialogs are created using the <dialogstart> tag**
  - You pass the URL of the document that you want to run.
- **Dialogs can be ended using the <dialogterminate> tag.**
  - This allows CCXML to end a dialog based on a external event such as someone calling you on a second line.
- **Dialogs can return data back to the CCXML platform**
  - In VoiceXML this is done with the VoiceXML <exit namelist=“a b c”/> tag.
  - This is exposed in the CCXML dialog.exit event.

CCXML <dialogstart> and VoIP

- **Uses the SIP INVITE message**
- **Passes the URL of the VoiceXML document inside the sip URL as defined in the IETF Internet Draft located at:**
  - http://search.ietf.org/internet-drafts/draft-rosenberg-sip-vxml-00.txt
- **Additional data can be passed in the body of the SIP message.**
  - VoiceXML variables
  - Platform Specific data
  - Billing information
- **The VoiceXML platform sees this as a new incoming call.**
  - Only minor modifications required to a SIP based platform to support this.

CCXML <dialogterminate> and VoIP

- **Terminates the currently running dialog on the VoiceXML Server.**
- **Uses the SIP BYE message**
- **Additional data can be passed in the body of the SIP message.**
  - Platform Specific data
  - Billing information
- **The VoiceXML platform sees this as a call disconnect.**
  - Only minor modifications required to a SIP based platform to support this.

VoiceXML <exit> and VoIP

- **Terminates the currently running dialog on the VoiceXML Server.**
- **Uses the SIP BYE message**
- **The values of the namelist vars are returned in the body of the SIP BYE message.**
- **Additional data can be passed in the body of the SIP message.**
  - Platform Specific data
  - Billing information
- **The CCXML platform sees this and returns a dialog.exit event to the CCXML document.**

Where to learn more.

- **Voice Browser Working Group**
  - http://www.w3.org/Voice
- **CCXML Specification**
  - http://www.w3.org/TR/2002/WD-ccxml-20020221
- **Voxeo Developer Community**
  - http://community.voxeo.com

Voice Browser Interoperation

Seamless Interconnection of Voice Applications

Voice Browser Interoperation Overview

- **Call:**
  - Audio path, signaling path, state information
- **Call Site:**
  - voice application that receives and makes telephone calls
- **Transfer *call* from one call site to another**
  - Messages, message sequence, and message data
  - Voice path setup and tear down
- **Transfer *data* from one call site to another**
  - Information about a user (name, customer ID, caller ID, etc.)
  - Originating call site information (e.g. "Acme Airlines reservations call center")
  - Application-specific information (e.g. travel itinerary)

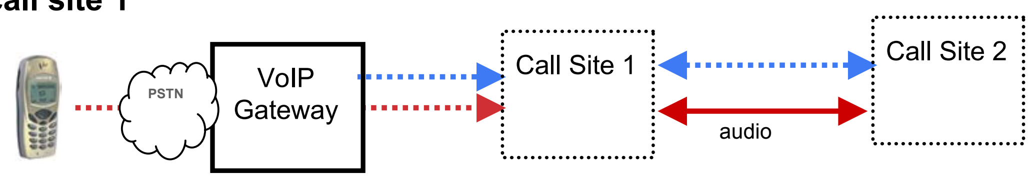
- **Transfer hotword listener request**
  - Downstream call site requests "browser"/voice dialer to listen for a site-specific hotword (e.g. "go to Acme Airlines")

Call Transfer between 2 sites

- **“traditional” VoiceXML `<transfer>`**
- **May share data via PSTN connection, e.g. ISDN UUI**
  - ***Note: AAI on `<transfer>` defined in VoiceXML 2.0***
- **May share data via IP network, e.g. SIP**
- **Audio: TDM or RTP**

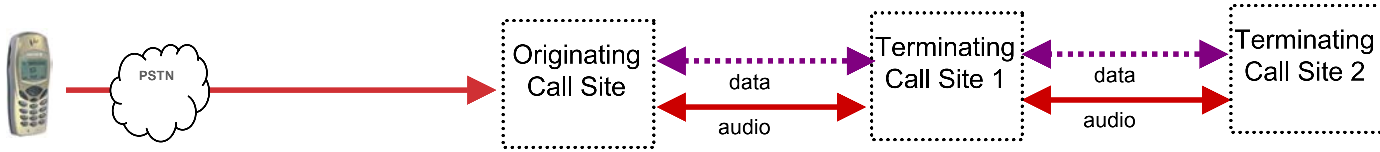
TDM Call hairpins through call site 1

VoIP Call hairpins through call site 1

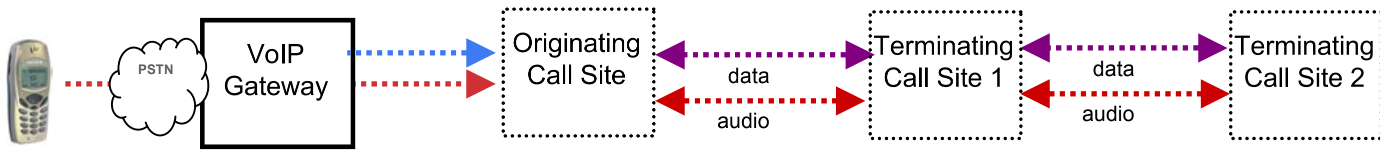


Call Transfer between 3 sites

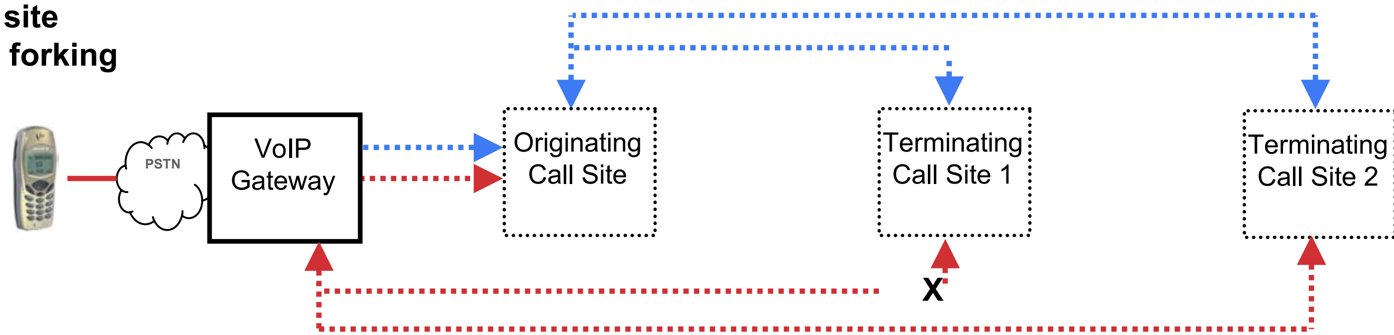
TDM Call hairpins through each call site



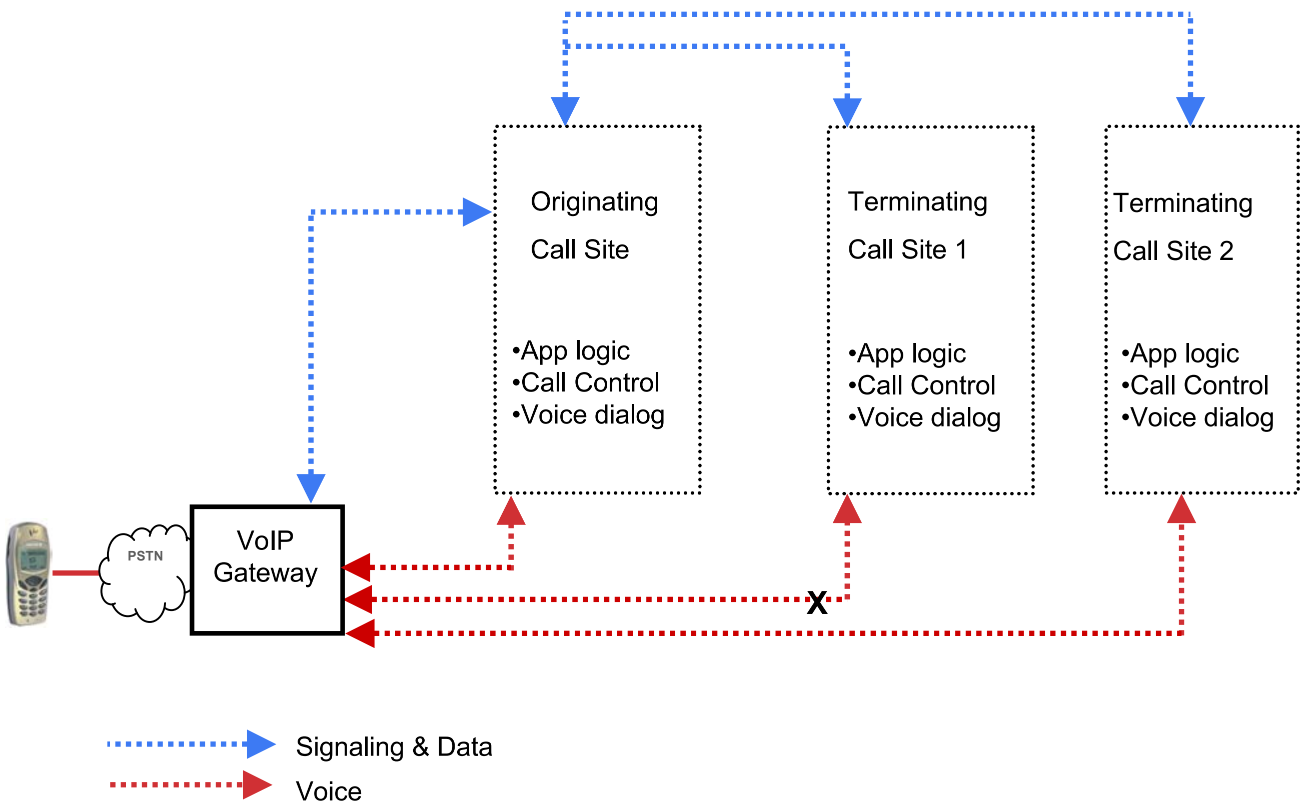
VoIP Call hairpins through each call site



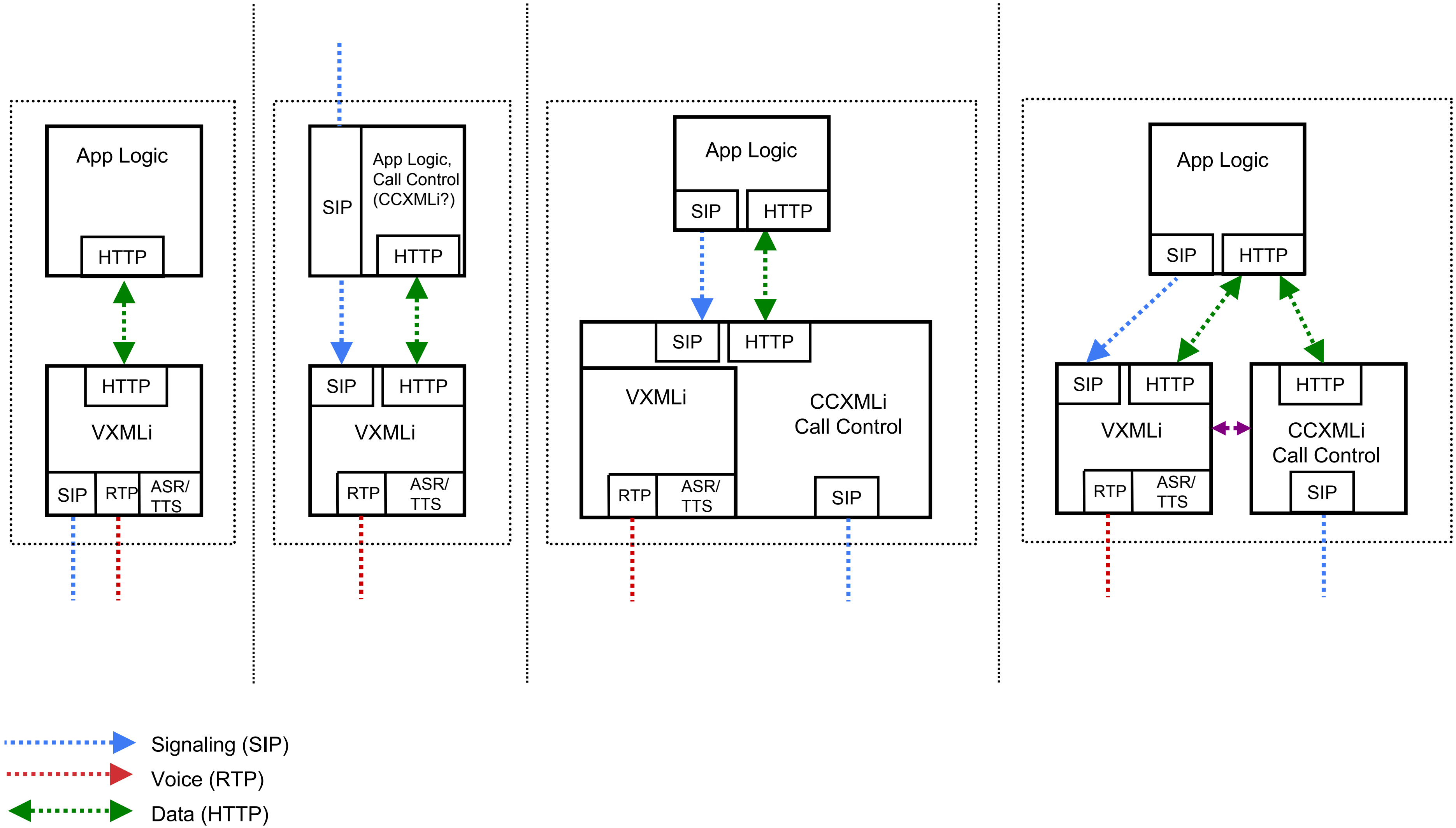
VoIP Call redirected to each call site using audio forking at gateway



Call Transfer between 3 sites

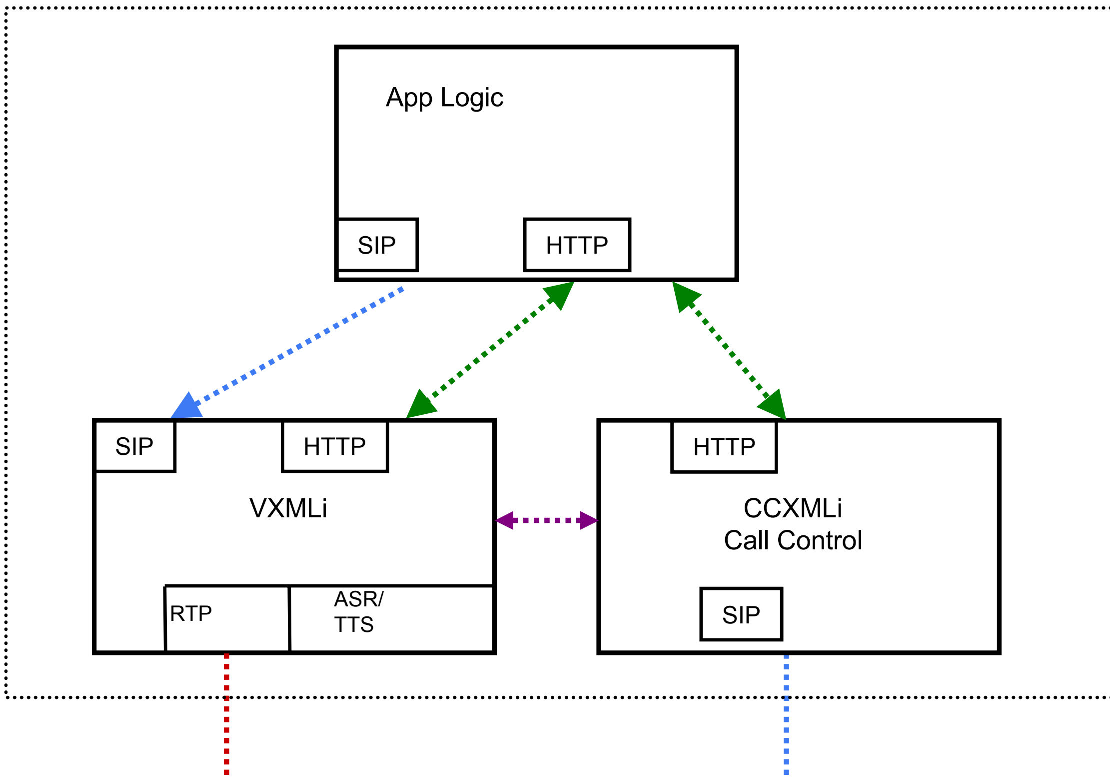


"Call Site" application architecture variations



CCXML-VoiceXML Interoperation

- **Communication between CCXMLi and VoiceXMLi**
- **Enhancements required in VoiceXML:**
  - Asynchronous external event handling
  - Unified *Connection Object* for connection data representation



Advantages of VoIP for VoiceXML Deployments

- Flexible Network Topology
- Simplified integration of voice dialog resources
- Vendor independence for network elements
- Separation of concerns: voice dialog resources vs. call control

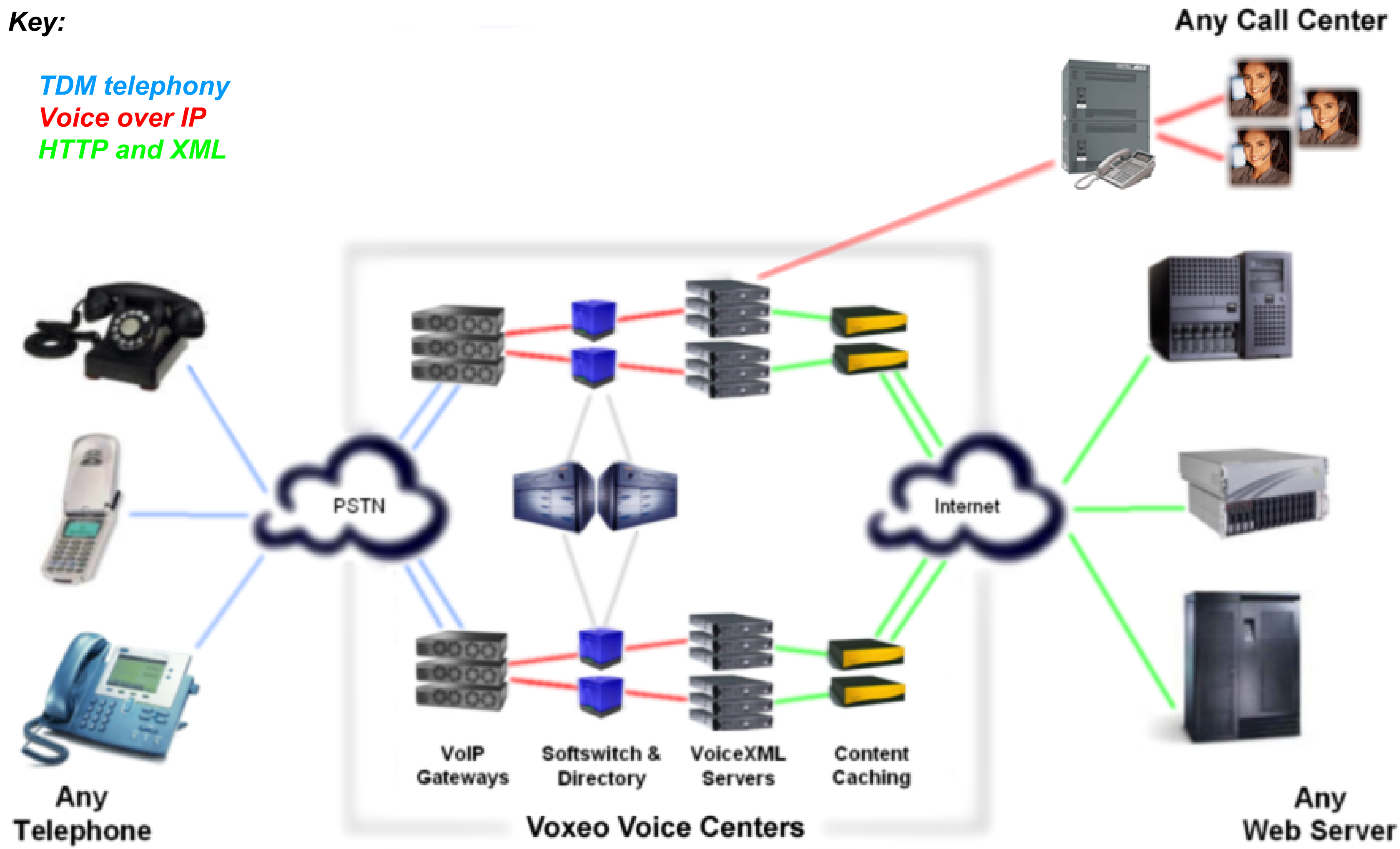
Voxeo.Net

A Voice over IP enabled managed VoiceXML Solution.

Voxeo Overview

- First implementation of the CCXML standard
- Is the only network built entirely with flexible VoIP technologies
- Was designed from day one to deliver managed voice hosting
- Features pre-deployed Voice Centers for unrivaled redundancy
- Includes leading call control and call center integration abilities
- Gives complete control via XML data and Web Services
- Is a certified Cisco Powered Network and Cisco TASP
- Builds on 7 years experience hosting Phone & Web applications

Voxeo System Architecture



voxeo

http://www.voxeo.com

http://community.voxeo.com

mailto:rj@voxeo.com

Nuance Communications

Software for a voice-driven world.

VoiceXML Planet Conference   San Jose, California USA   2002 March 21

Nuance Overview

- **VoiceXML industry leadership:**
  - Two editors of W3C VoiceXML 2.0 specification
  - Chair, W3C Voice Browser Interoperation Subgroup
  - Editor of W3C Speech Synthesis Markup specification
  - W3C Speech Recognition Specification, Call Control, Natural Language Semantics, etc.
  - Chair, VoiceXML Forum Conformance Committee
- **Nuance Voice Web Server**
  - Complete VoiceXML 2.0 solution
- **Nuance 8 ASR**
  - SayAnything™ Natural Language
  - Just-in-Time Grammar Compliation
  - Best performance in noisy conditions
- **Nuance Vocalizer**
  - Industry's most natural speech synthesis
- **Nuance Verifier**
  - Most accurate speaker verification

V-World

Nuance Speech Conference

Orlando, Florida USA
April 29 – May 2, 2002

http://www.nuance.com

http://extranet.nuance.com

mailto:ken@nuance.com