

# VoiceServer 2.0 产品白皮书



声软网络技术(中国)有限公司

2002.9

## 目 录

1.	概述	1
1.1.	市场需求	1
1.2.	技术基础	1
1.3.	市场机遇	2
1.4.	小结	2
2.	VoiceXML 的优势	3
2.1.	对企业用户	3
2.2.	对服务提供商	3
2.3.	对软件开发商	3
3.	VoiceServer	4
3.1.	概述	4
3.2.	产品特点	4
3.2.1.	技术领先	4
3.2.2.	功能独特	4
3.2.3.	开放结构	4
3.2.4.	高性能	5
3.2.5.	高可靠性	5
3.2.6.	管理维护	5
3.3.	产品功能	5
3.3.1.	语音识别	5
3.3.2.	语音合成	5
3.3.3.	收发传真	5
3.3.4.	呼入排队	5
3.3.5.	电话呼出	5
3.3.6.	呼叫转移	6
3.3.7.	录制语音	6
3.3.8.	播放语音	6
3.3.9.	接收按键	6
3.3.10.	业务详单	6
3.3.11.	Javascript	6
3.3.12.	流程控制	6
3.3.13.	业务组件	6
3.3.14.	平台管理	6
3.4.	系统结构	7
3.4.1.	系统结构图	7
3.4.2.	语音浏览器	7
3.4.3.	服务适配层	8
3.4.4.	服务提供层	8

3.4.5.	应用服务器.....	8
3.4.6.	业务访问层.....	8
3.4.7.	操作、管理和维护 (OA&M) .....	8
4.	语音应用 .....	9

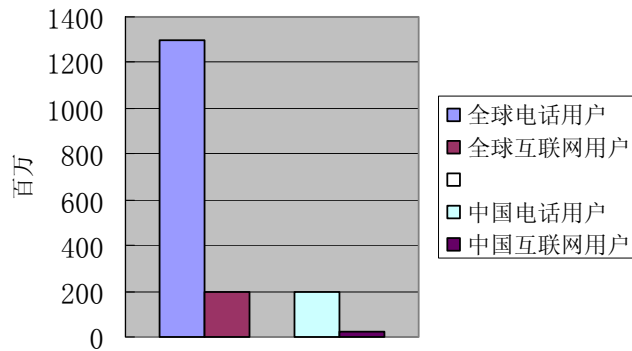
## 1. 概述

### 1.1. 市场需求

据权威机构预测，到 2005 年全球语音应用产值将达到 450 亿美元。语音应用之所以成为新一轮的市场焦点，是因为：

- 电话尤其是移动电话的日益普及
- 人们更需要及时、方便地得到信息
- 企业、服务提供商需要提供更完善的语音服务。

悠久的历史使电话无疑是最普及的通讯工具，其操作简易性更是老少皆宜。如下图所示，尽管互联网方兴未艾，然而在全球范围内，电话用户的数量仍数倍于互联网用户，在中国更是达到 10 倍以上。特别是，近年来移动通讯发展迅猛，移动电话的销售量已超过汽车和 PC 机的总和，预计 2003 年末全球移动用户数将突破 10 亿。



信息技术革命将人类社会推进到信息社会，企业的经营管理、个人的休闲理财都离不开信息。人们比以前的任何时期都更加渴望能够随时随地得到信息。而语音作为人类最原始、最自然的交流方式，最受人们的青睐。Evans Group Research 对 250 个使用不同语音系统的用户进行的调查表明，83% 的用户更喜欢使用语音系统而不是按键式系统。

在激烈的市场竞争中，企业、服务提供商一直在寻求完善客户服务的途径。金融、电信等领域实施呼叫中心的成功经验，使企业、服务提供商体会到语音服务带来的经济效益和社会效益。提供更为人性化、智能化的语音服务系统，提供 24 小时全天候的自助服务系统，已成为企业和服务提供商重要的市场策略。

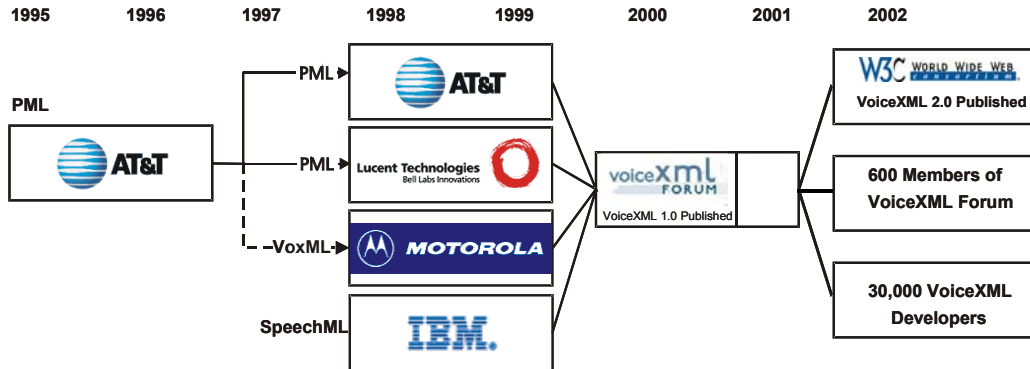
### 1.2. 技术基础

近几年来，计算机语音识别（ASR）和语音合成（TTS）技术的突破给语音应用提供了坚实的基础，与此同时 VoiceXML 的标准化更使语音应用生机勃勃。

语音识别技术是指将人说话的音频信号转换为可被计算机所识别的文字信息，而语音合成技术则是指将文字信息转变为音频信息的技术。事实上，语音识别产品进入大众视野只有短短的 4 年。1998 年 IBM 率先推出 ViaVoice 引发了语音应用的热潮，研究机构和企业厂家纷纷研制各具特色的语音产品。通过识别算法的改良，同时得益于硬件能力的提高，今天优秀的语音识别系统的识别率高达 95% 以上，完全能够满足商业应用的需求。而语音合成技术则已经能够实现自然、流畅的文字

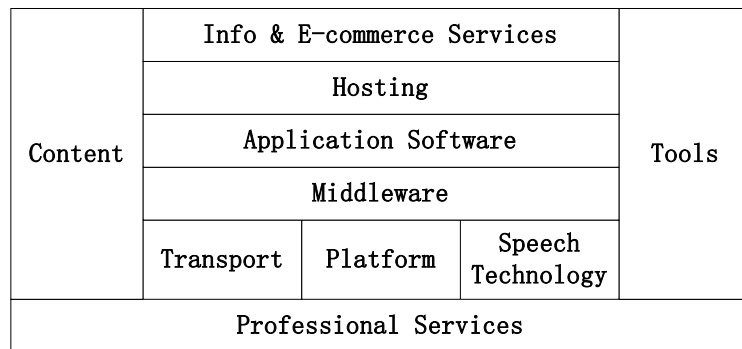
合成。

VoiceXML 是由 IBM、Lucent、Motorola 和 AT&T 四家公司于 2000 年提出的一种应用于语音浏览的标记语言。如下图所示，VoiceXML 是各厂商多年经验积累的产物，VoiceXML Forum 更是得到业界广泛的支持，同时 VoiceXML 由 W3C 组织标准化，成为事实上的工业标准。



### 1.3. 市场机遇

语音应用市场的构成如下图所示：



一般地：

- 企业用户可以在其电子商务或呼叫中心平台的基础上提供语音访问界面，以为最终用户提供更为体贴、便利的服务；
- 服务提供商可以提供语音应用托管业务，或直接提供各具特色的语音应用；
- 应用软件开发商、系统集成商可以在应用中集成语音功能，为用户提供语音访问界面；
- 语音应用专业 IT 厂商则提供应用软件 (Application Software)层以下以及工具(Tools)部分，为各类语音应用提供坚实的平台以及开发、管理和维护环境。

### 1.4. 小结

电话是人类最普及的通讯工具，语音是人类最自然的交流方式。电话依旧是用户向企业、服务提供商寻求服务的最主要方式之一。

计算机语音识别、语音合成以及语音浏览技术的涌现，掀起自图形界面之后新一轮的人机界面革命，实现了人类通过对话与计算机进行交互的梦想。人们可以通过电话随时获得最新的信息和体贴的服务。

企业、服务提供商应该把握市场机遇，使自身在激烈的市场竞争中脱颖而出。

## 2. VoiceXML 的优势

### 2.1. 对企业用户

- 电话是最终用户联系企业的主要联系方式，VoiceXML 以其简洁易用性，为企业提供语音门户，使企业可以提供更为丰富的 24 小时客户服务、信息查询、在线交易等自助服务，树立企业品牌形象，产生更为可观的经济效益和社会效益；
- VoiceXML 平台可以作为低成本的呼叫中心，或者作为以 Web 为中心的电子商务平台的补充，使企业在充分保护现有投资的基础上，进一步开拓销售渠道、完善客户服务、提高运作效率、降低运作成本；
- VoiceXML 符合国际标准，建立在 XML 和 Web 技术的基础上，使得企业用户可以摆脱对专用语音系统厂商的依赖，利用企业自身技术力量，特别是利用企业对自身客户需求的深刻理解，在最短的时间内推出满足市场需求、符合企业发展目标的语音应用；
- 传统的语音应用主要以用户呼入为主，VoiceXML 由于其流程控制的简易性，使企业可以方便地建立各类用户回访、电话确认、电话通知应用，从而更好地体现企业“用户至上”的宗旨。

### 2.2. 对服务提供商

- 及时推出满足用户需求的服务是服务提供商在市场中赢得先机的关键，采用 VoiceXML 可以大大缩短应用开发周期，同时服务提供商完全可以依托自身的技术力量，推出独具特色的新业务；
- 服务提供商可以依托自身丰富的资源给中小型企业托管语音应用，VoiceXML 以其简洁性，大大降低了服务提供商项目管理、系统维护的支出，从而降低运营成本；
- 由于 VoiceXML 基于通用技术，相比专用 IVR 系统，人员培训、系统维护等费用将有所减少，同时语音应用可以运行在任何支持 VoiceXML 的语音平台。

### 2.3. 对软件开发商

- 可以缩短语音应用开发周期。与传统的 IVR 系统不同，即使在硬件设备尚未就绪的情况下，也可以在集成开发环境中开发、调试业务应用，从而缩短项目进度；据估计，与传统的 IVR 开发相比，业务流程编码时间可以缩短 12 倍，而整个项目开发进度可以缩短一半；
- 可以降低服务开支。由于传统的 IVR 系统往往是专用的，专业人员的培训、用户培训等服务性开支较大，而 VoiceXML 作为国际标准，特别是基于 XML 和 Web 等通用技术，服务性开支减少；
- 可以减少软件维护、升级的工作量。在整个产品的生命周期中，维护工作将占很大的比例；作为一种脚本语音，VoiceXML 的维护工作量大大减少，而且配置管理工作也相对简单，从而从整体上降低了项目管理的复杂度，以节省软件开发成本、提高软件发布效率。

## 3. VoiceServer

### 3.1. 概述

VoiceServer 是公司推出的新一代产品，主要面向运营商、服务提供商及企业用户，提供基于 VoiceXML、支持语音识别/合成的语音服务系统。同时作为技术平台提供商，公司提供专业的业务咨询、项目管理、人员培训等综合服务。

VoiceServer 支持与内容服务平台或电子商务平台的集成，向最终用户提供电话服务途径；支持呼入排队、转人工台功能或与现有呼叫中心平台的集成。VoiceServer 可以作为服务提供商、企业的语音门户，向最终用户提供 7x24 小时的个性化自助服务。

### 3.2. 产品特点

#### 3.2.1. 技术领先

- 自主产权，软件开发工作完全在中国大陆进行，能为客户提供最为全面的本地化服务，按客户需求快捷、经济地定制业务
- 设计灵活，可以作为语音门户，或作为低成本的呼叫中心实施
- 采用开放技术，能够与 .NET、J2EE 或数据库实现的业务逻辑接口，使客户无需很大投资即可在现有系统的基础上引入语音服务
- 支持 VoiceXML 2.0 规范，符合国际规范，业务逻辑独立于平台，使客户可以自由地选择平台提供商
- 具备丰富的 VoiceXML 开发、调试环境，能高效、快捷地实现满足客户需求地业务逻辑
- 外部数据存储和交换采用 XML 格式，能与客户现有地其他信息系统方便地交互

#### 3.2.2. 功能独特

- 支持呼入排队和呼叫转移功能，使企业能够以最小的投资实现基本的呼叫中心功能
- 支持与呼叫中心 CTI 服务器接口功能，对已有中、大型呼叫中心的客户，能供通过采用该平台提供更为完善的自服务功能
- 支持呼出功能，在呼出成功后执行 VoiceXML 脚本，使服务提供商、企业提供更为主动的客户服务成为可能
- 支持传真功能，使服务提供商与企业用户可以通过传真沟通产品信息、反馈服务意见等
- 支持电话会议，使服务提供商可以提供如电话 ICQ 功能等

#### 3.2.3. 开放结构

- 采用面向对象分析和设计，具备良好的可扩充性和可维护性，使客户可以根据业务量需求平滑地升级系统
- 通过引入服务适配层将核心逻辑独立于硬件设备和第三方引擎，能方便地和不同地硬件设备、第三方引擎集成
- 支持各种组件技术、通讯技术实现与后台业务平台的通讯，最大程度地复用客户现有地信息资源，大大地减少投资

#### 3.2.4. 高性能

- 采用 C++编程
- 采用异步模式提供实时服务
- 采用多线程技术并发处理

#### 3.2.5. 高可靠性

- 提供负载平衡的资源管理，能提供不间断的服务
- 提供信令、资源的备份方案

#### 3.2.6. 管理维护

- 采用 SNMP 协议
- 提供远程、本地的实时监控功能
- 支持在线更新业务
- 提供完整的报表统计功能。

### 3.3. 产品功能

#### 3.3.1. 语音识别

可选功能。

系统支持语音识别功能，并且系统能同时支持语音输入和 DTMF 键输入。

#### 3.3.2. 语音合成

可选功能。

系统支持语音合成功能，支持中文普通话、中文粤语、英语等。语音菜单、语音提示均可以通过语音合成后播放。

#### 3.3.3. 收发传真

可选功能。

系统支持传真功能，用户可以通过该系统接收、发送 TIFF 文件。

#### 3.3.4. 呼入排队

可选功能。

系统支持按话务员组、技能组等路由的呼入排队功能。

#### 3.3.5. 电话呼出

可选功能。

系统支持呼出功能，其呼出任务管理系统支持呼出的重试、超时等功能。



### 3.3.6. 呼叫转移

可选功能。

系统支持呼叫转移功能，提供呼叫转移、电话会议等功能。

### 3.3.7. 录制语音

基本功能。

系统支持录音功能，录音内容可以回放、保存等。

### 3.3.8. 播放语音

基本功能。

系统可以播放预先录制的语音文件，支持的语音文件格式包括 PCM linear、PCM a-law、PCM u-law、ADPCM 等。

### 3.3.9. 接收按键

基本功能。

系统可以接收用户可以通过 DTMF 键输入的信息或选择的菜单。

### 3.3.10. 业务详单

基本功能。

系统产生 XML 格式的详细话单供帐务系统处理。

### 3.3.11. Javascript

基本功能。

系统支持在 VoiceXML 中的 javascript 及标准的 Session 变量，支持的预定义类型包括数字等。

### 3.3.12. 流程控制

基本功能。

系统提供 javascript 实现的 VoiceXML 流程控制模板，以提供编写结构化 VoiceXML 的方式。

### 3.3.13. 业务组件

可选功能。

系统支持在 VoiceXML 中使用客户定制的业务组件。

### 3.3.14. 平台管理

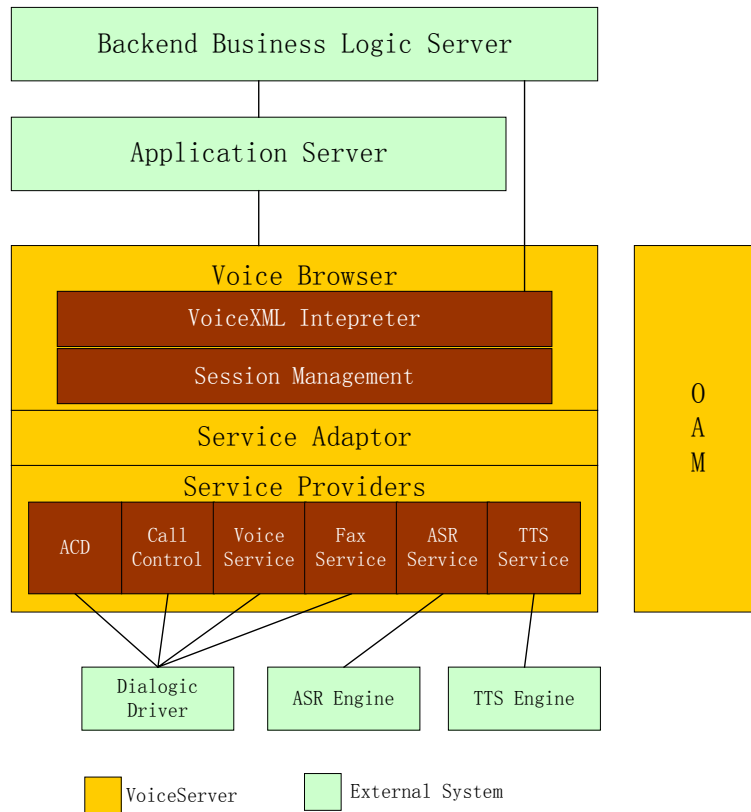
基本功能。

系统提供平台管理系统，提供业务统计、业务监控功能。

### 3.4. 系统结构

#### 3.4.1. 系统结构图

VoiceServer 系统结构如下图所示：



#### 3.4.2. 语音浏览器

VoiceXML 语音浏览器（Voice Browser）是系统的控制中心，包括 VoiceXML 解析和会话控制两部分。其中，VoiceXML 解析部分功能如下：

- 通过文件系统或 http 获取 VoiceXML 脚本
- 采用 XML DOM 解析程序解析脚本
- 根据 VoiceXML 规范中定义的 Form 解释算法解释执行脚本
- 根据不同标记（tag），或执行控制流，或执行 javascript 脚本，或产生相应的控制命令。

会话控制部分的功能如下：

- 根据 VoiceXML 解析器的控制命令按要求执行其同步、异步的操作
- 驱动服务提供层完成呼叫控制、语音录放、语音识别、语音合成等操作
- 协调各服务提供者间的时序
- 向 VoiceXML 解析器返回操作结果。

### 3.4.3. 服务适配层

服务适配层(Service Adaptor Layer)定义服务提供者的接口。通过提供不同的服务提供者，系统可以方便地支持不同的硬件设备，支持不同的第三方语音识别、语音合成引擎。

系统定义的服务提供者接口包括：

服务接口	功能
呼叫控制服务	提供呼入、呼出、挂机、呼叫转移、会议电话的控制
呼叫排队服务	提供呼叫排队、呼叫路由功能
基本语音服务	提供语音播放、DTMF 接收、语音录制功能
传真服务	提供接收传真、发送传真功能
语音识别服务	提供语音识别功能
语音合成服务	提供语音合成功能

### 3.4.4. 服务提供层

服务提供层(Service Provider Layer)实现服务适配层的接口。系统目前提供的服务提供者包括：

服务接口	服务实现
呼叫控制服务	Dialogic 模拟、数字信令
基本语音服务	Dialogic 语音功能
传真服务	Dialogic 传真功能
语音识别服务	Nuance 语音识别引擎
语音合成服务	L&H 语音合成引擎

### 3.4.5. 应用服务器

应用服务器(Application Server)用于存放或产生 VoiceXML 脚本，可以是文件系统或 Web 服务器。一般的，文件系统应用于静态脚本，Web 服务器应用于动态脚本。对于动态脚本，ASP、JSP、PHP 等任何用于动态产生 HTML 脚本的技术均适用于 VoiceXML。对于动态脚本，应用服务器还负责与后台业务逻辑服务器接口。

### 3.4.6. 业务访问层

业务访问层(Business Access Layer)用于访问企业、服务提供商的后台业务逻辑，系统支持服务器端和客户端两种方式。

服务器端方式一般用于动态 VoiceXML 脚本，在 ASP/JSP/PHP 中访问后台业务逻辑，并根据业务操作结果向语音浏览器返回 VoiceXML 脚本。

客户端方式利用 VoiceXML 的<object>标记，在语音浏览器中执行业务操作。系统支持将 DLL、COM/COM+组件、JavaBean、xml/http 等作为<object>访问后台业务逻辑。

### 3.4.7. 操作、管理和维护 (OA&M)

操作管理维护终端提供业务监控、业务管理和业务统计功能。

## 4. 语音应用

VoiceServer 具有广泛的应用前景，可以应用于：

- 信息查询
- 自助服务
- 通知服务
- 消息服务。

部分应用举例如下表所示：

应用类型	应用举例
信息查询	航班查询、股市信息、路线信息、公共事业
自助服务	银行业、证券业、电信业以及呼叫中心、电子商务的自助服务
通知服务	日程安排通知、股票到价通知、缴款通知等
消息服务	邮件服务、传真服务、语音信箱、统一消息服务等