



Microsoft®
Speech Server 2004

Product Datasheet

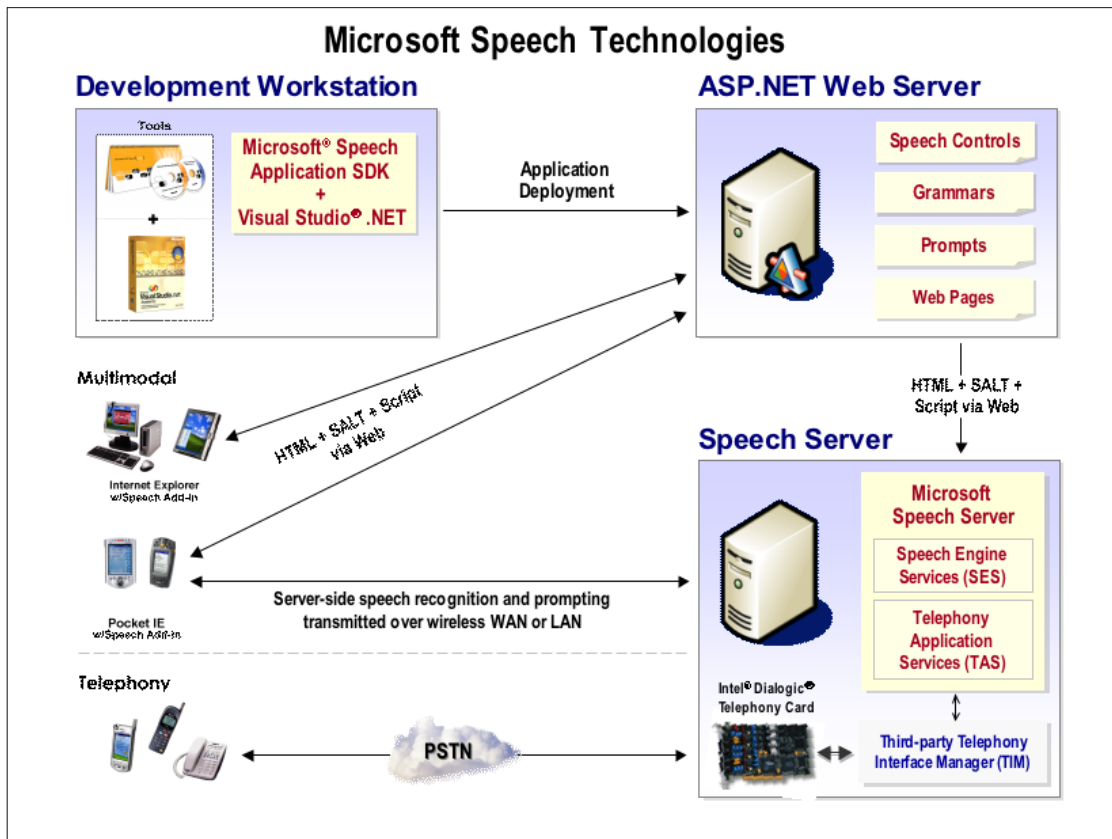
Microsoft® Speech Server 2004 (MSS) combines Web technologies, speech-processing services, and telephony capabilities into a single, integrated system. The server enables companies to unify their Web and telephony infrastructure, and extend existing or new ASP.NET Web applications for speech-enabled access from telephones, mobile phones, Pocket PCs, and Smartphones. Microsoft Speech Server is a flexible and integrated speech platform that dramatically reduces the complexity and cost of developing and deploying speech applications.

ASP.NET applications from a broad variety of industries can be speech-enabled using Microsoft Speech Server. Examples include contact center self-service applications such as call routing, and customer account and personal information access. Many other contact center

speech-enabled applications are possible including travel reservations, financial and stock applications, and CRM deployments. Information Technology (IT) groups can benefit from speech-enabled applications in the areas of sales and field-service automation, e-commerce, auto-attendants, helpdesk password reset applications, and speech-enabled network management, to name just a few.

Companies can use Microsoft Speech Server 2004 to deploy telephony-only DTMF (touch-tone keypad) applications, speech-enabled telephony applications, and multimodal (mixed speech/visual) applications today, all while leveraging their existing IT and contact center technology infrastructure and assets.

The diagram below represents Microsoft Speech Server and its components: Speech Engine Services (SES), Telephony Application Services (TAS), and the Speech Application SDK



Telephony Application Services (TAS)

The TAS server component provides a scalable and fault tolerant SALT application runtime environment for telephony applications.

Core Functionality:

- Supports multiple concurrent speech applications running on a single TAS node
- Processes application dialogues, call control, speech requests, and call logging
- Interprets SALT-enabled HTML via SALT interpreter
- Handles DTMF and speech applications
- Coordinates all media processing and speech recognition requests

Telephony Functionality:

- Supports a wide range of call control features (call answer, call placement, call transfer etc.) through Intel or Intervoice Telephony Interface Managers (TIMs)
- Handles up to 96 telephony ports of continuous, fully loaded speech recognition per TAS node
- Unlimited scale-out by adding more nodes
- Barge-in support
- Analog telephony channel support: 4-port and 16-port channels supported
- Digital telephony channel support: 24, 48 and 96 channels supported
- CAS (Channel Associated Signalling) support
- CCS (Common Channel Signalling) support

Standards Support:

- ECMA-323 (CSTA) XML for call control
- SALT (Speech Application Language Tags)
- T1 ISDN PRI
- T1 CAS (FXS)

Speech Engine Services (SES)

The SES component provides scalable and fault-tolerant access to text-to-speech and speech recognition resources for telephony and multimodal applications.

Core Functionality:

- Enables scale-out and multiple-node redundancy using standard IP load balancing
- Supports dynamic grammar loading and caching, and compiled grammars
- Allows mixed-initiative applications
- Provides robust data logging for application dialog and speech recognition engine tuning

Speech Recognition Engine:

- Includes Microsoft enterprise-grade ASR (Automatic Speech Recognition) engine
- Supports Scansoft OSR (available separately)

Text-to-Speech Engine:

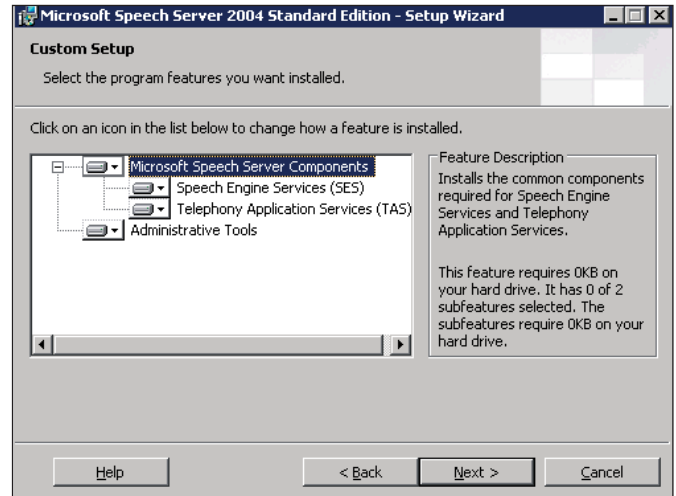
- Includes Scansoft Speechify TTS engine out-of-the box

Prompt Engine:

- Includes Microsoft prompt engine for playing pre-recorded .WAV prompts

Standards Support:

- W3C SRGS grammars
- W3C SSML



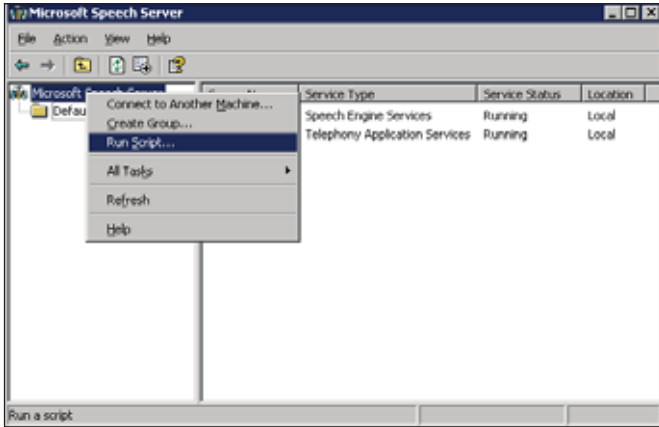
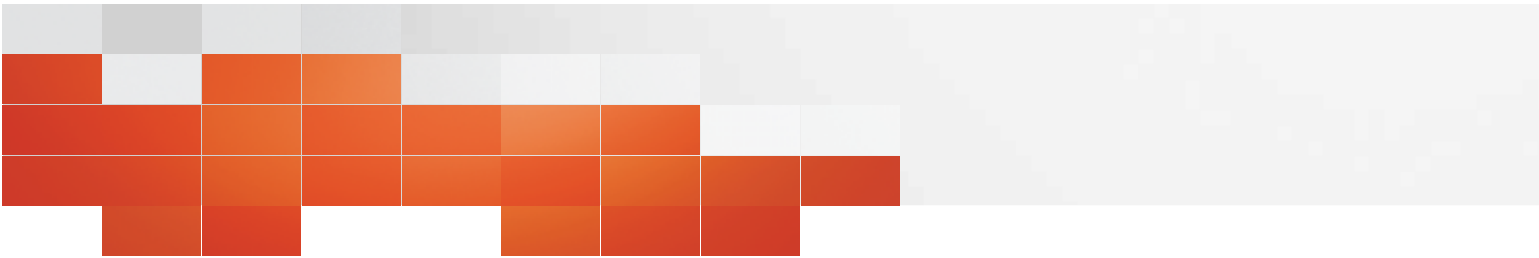
Speech Server 2004 installs just like other familiar Microsoft server products

Microsoft Speech Server Administration and Manageability Features

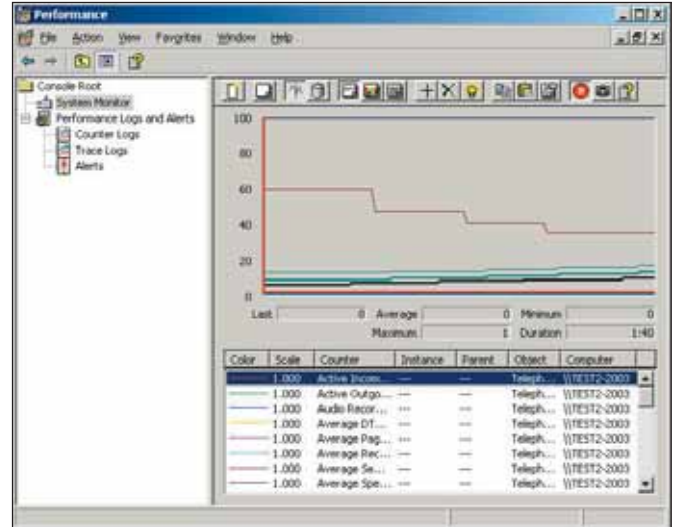
Microsoft Speech Server provides functionality that enables administrators to easily configure the core functionality required to cost-effectively manage the server and speech applications. Administration features support both local and remote administration.

Standard Microsoft Window Server System™ features leveraged by MSS include:

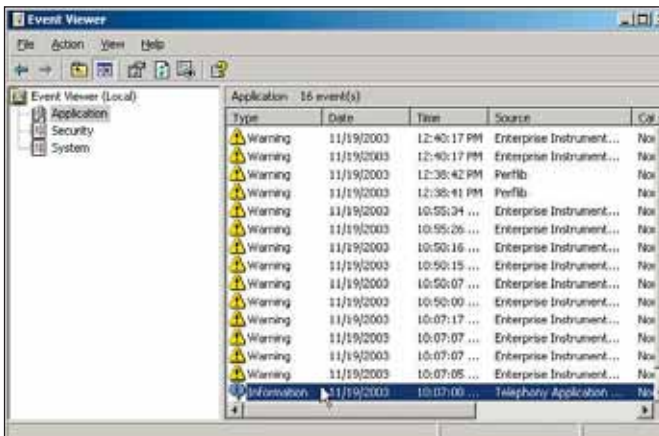
- Microsoft Management Console support (snap-ins for SES, TAS configuration, and app deployment included)
- Windows® Script Host for Windows Management Information enables operations automation via scripts
- Configuration, back up, and optimization (scheduled or unscheduled)
- Windows event logging and event tracing
- Performance Monitor
- System health monitoring



Run scripts from within MMC to tailor a unique deployment



Quickly assess the impact of upgrades on performance using counters, tracing, and alerts available in the Performance Monitor



Use the Event Viewer to quickly find and fix errors or warnings

Key MSS administration features and functionality include:

- Application configuration hot swap updates without requiring system restart
- Deployment wizard for rapid deployment of speech applications
- Ability to add new SES and TAS servers and replicate settings across multiple servers
- Easily define and manage application start pages, memory limits, time-out parameters, tuning parameters
- Define URL for SES server and optimal number of ASR engine instances
- Set/alter location of recording directory
- Administer barge-in settings
- Resource administration: audio connection IP, audio connection port, shutdown time-out, maximum cache size, cache location
- Rich and detailed logging for application troubleshooting and tuning
- Set a scheduled TAS process recycle

General system operation features include:

- Back up of SES and TAS settings
- Event and performance monitoring
- Debugging tools
- Trace logs and Log Analysis Tools
- GUI tool for managing Network Load Balancing

Reliability Features

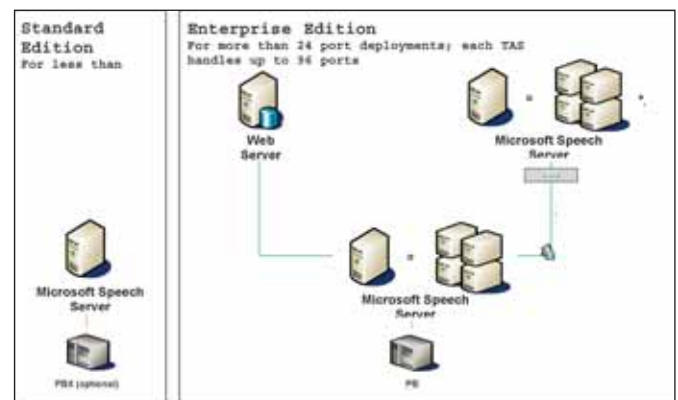
Speech Server has been architected to provide a highly reliable and available server solution for speech application deployments and was designed to deliver the following features and functionality to assure high availability:

- Robust fault-tolerant architecture to provide maximum availability
- Fail-over within a node provided by redundant server processes
- Stateless servers plus hardware/software load balancing drive high system availability
- Telephony Application Services
 - Each SALT interpreter and call interaction is isolated
 - Call extensibility handled out of process
 - Automatic SALT interpreter and worker process recycling
- Speech Engine Services
 - Node configuration supports all TTS or ASR deployments
 - Full featured grammar cache prevents uncommanded reboots
 - Supports load balancing of requests within a node
 - Hardware load balancing supported for server farm across nodes
- Supports Windows Server™ 2003 reliability features

Scalability Features

The ability to have a flexible scaling solution is designed into the architecture for each edition of Microsoft Speech Server. You can choose the edition that scales to meet your business needs today and tomorrow based on the following scalability features of MSS:

- Speech Server Standard Edition – deployments of up to 24 concurrent calls on one computer (TAS, SES, and Web Server together). Multiple Standard Edition computers may be connected to a telephony switch for scale requirements.
- Speech Server Enterprise Edition – unlimited scale-out with deployments of up to 96 ports per distributed TAS and SES servers. Multiple individual TAS and SES servers may be added for scale-out, with load balancing across TAS and SES servers.
- Telephony Application Services
 - Automatic adjustment of SALT interpreters based on load
- Speech Engine Services
 - Resource requests balanced within a node
 - Scale-out across nodes via a hardware load balancer



Microsoft Speech Server scalability



Security Features

Increased system security is now more critical than ever before. As part of the Microsoft commitment to reliable, secure, and dependable computing, the company has intensely reviewed Microsoft Speech Server 2004 to identify possible fail points and exploitable weaknesses. MSS provides many important new security features and improvements over existing industry security solutions including:

Network communication

- Enforces authentication/authorization on HTTP requests
- Provides secure audio channel between client and Speech Server
- SSL/https is supported if enabled

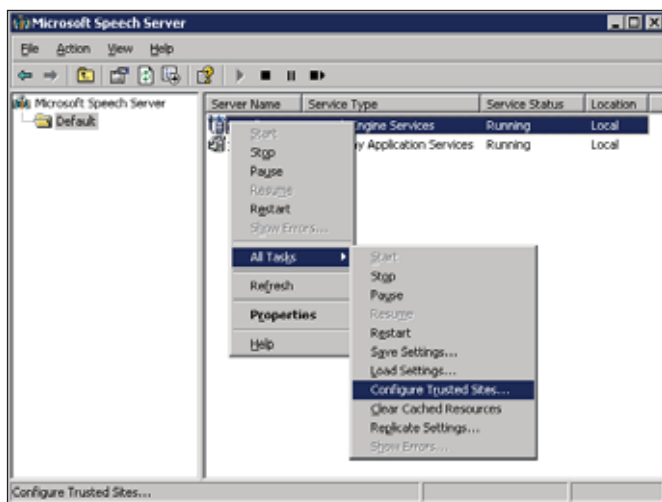
Application execution

- Resource loading limited to trusted sites
- Restricts access of server-side objects
- Ability to restrict logging of personal data

Service execution

- Services run under minimally-privileged accounts
- Security-appropriate ACLs on all resources
- Can restrict outbound dialing/transfers
- Provides auditing capabilities

Set access control via permissions, and configure trusted sites for content retrieval

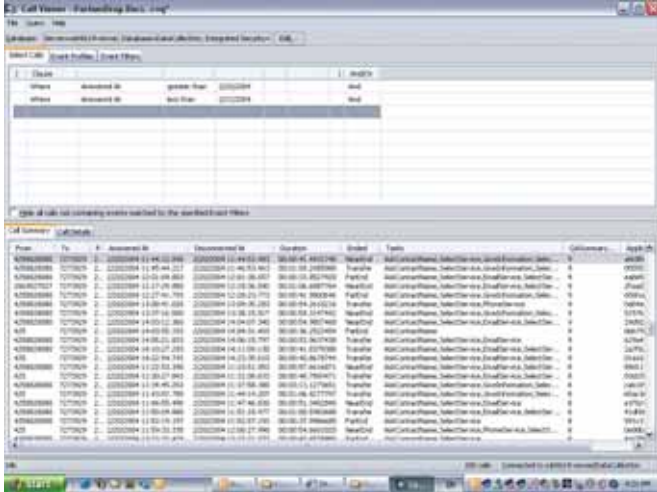


Easy to access and configure the Trusted Sites list

Reporting and Analysis

Built on Microsoft SQL Server 2000 Reporting Services, Microsoft Speech Server provides robust, out-of-the box reporting and analysis tools and pre-designed reports for both IT and Contact Center decision makers responsible for the business aspects of the speech solution. In addition, technology implementers will find a rich set of reporting and analysis tools for tracking, debugging, and tuning the deployed speech application. Custom reports can easily be created as well. Reporting updates can be scheduled or created on-demand. Key features for the reporting and analysis tools include:

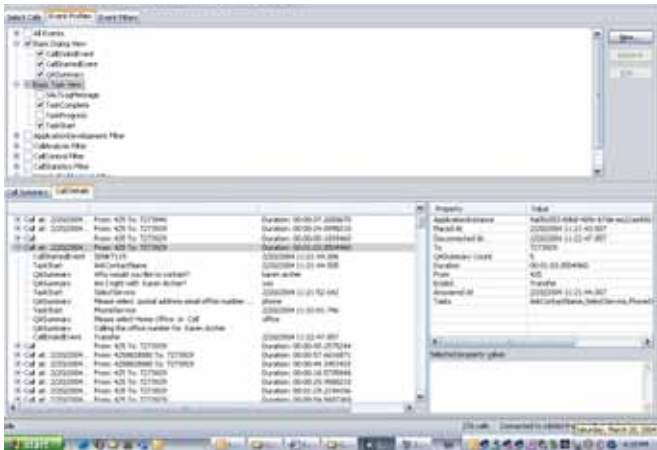
- Call Viewer – for developers and analysts needing detailed analysis of telephony calls and application information
 - Troubleshooting problematic calls
 - Analyze queries over times, events, and particular call properties
 - Retrieve all calls symptomatic of the problem
 - Drill down into detailed per call specifics
 - Flexible views, according to event filters selected
 - Conditional event filtering to see specifics
- Speech Application Reports – for business decision makers, IT and Contact Center Administrators to monitor the call flow and application usage from the business perspective
 - Built on SQL Server Reporting Services
 - A single, comprehensive reporting platform
 - Simplified report authoring
 - Flexible report management
 - Report delivery tailored to enterprises
 - Pre-designed, common enterprise reports included
 - System usage
 - Service completion
 - Performance
 - User behavior
 - Custom reports can easily be built for specific business needs



Use reports to track all calls coming into the system and their resolution



Easily generate reports to gain insight into your business and application usage



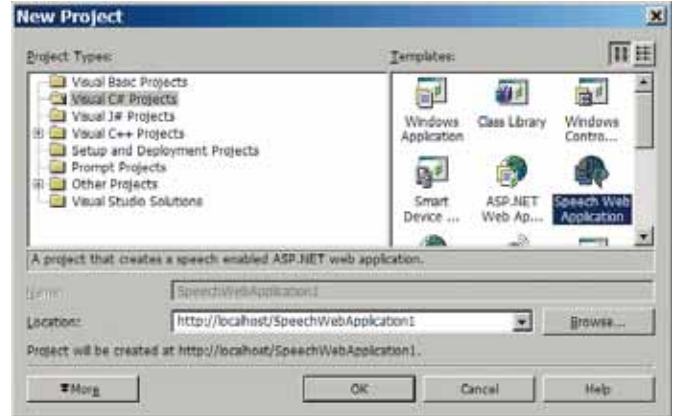
Drill into all call specific details for application tuning and analysis



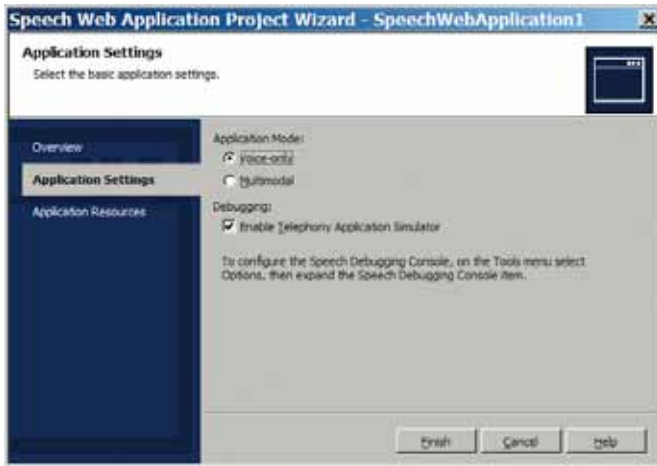
Speech Application Software Development Kit (SASDK)

The SASDK enables speech and Web developers to create and debug speech-enabled ASP.NET applications. The SASDK includes the following features:

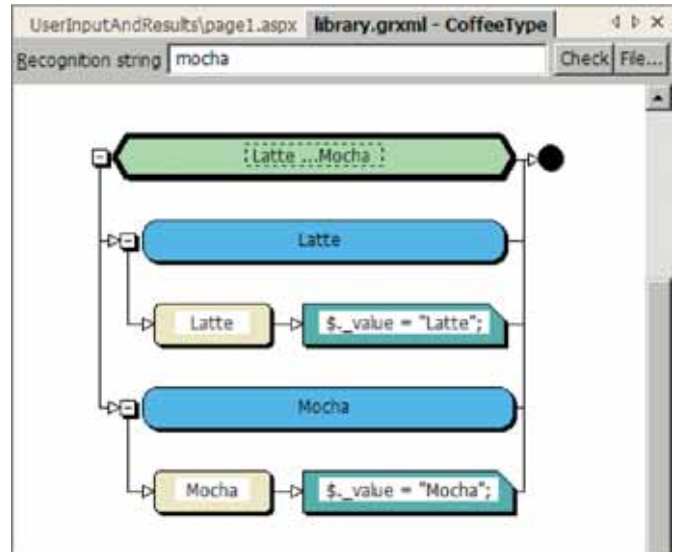
- Full integration with Visual Studio® .NET 2003
- Graphical grammar editor
- Prompt editor, management and alignment tool
- Debugging console
- ASP.NET speech controls
- Debugging and deployment wizards
- Dialog editor
- Application speech controls
- Speech control editor
- Speech Web application template
- Call inspection, log analysis, and reporting tools
- Telephony application simulator
- Tutorials and documentation
- Samples and reference applications



The New Project creation process for speech applications has the same look and feel as the creation process for other Visual Studio .NET projects



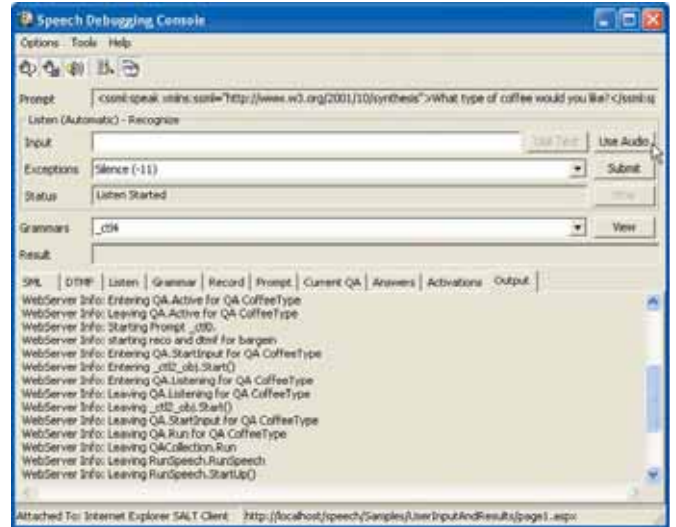
Start developing your applications quickly with the Project Wizard



The Grammar Editor lets you graphically design and modify your grammar

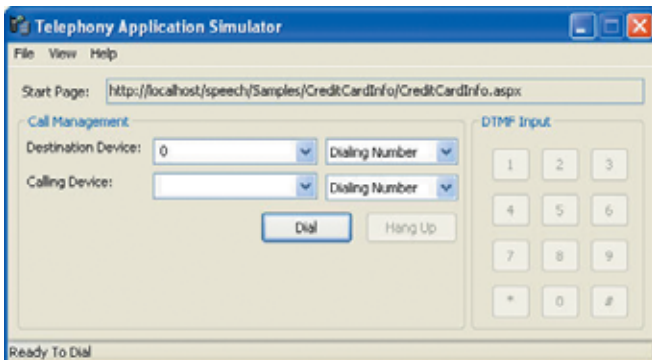
Transcription	Display Text	H...	H.	Wave
[You can say (BadContext...]	You can say Next, Previo...		✓	Rec0000.wav
[Next (NoContext)]	Next		✓	Rec0001.wav
[Previous (NoContext)]	Previous		✓	Rec0002.wav
[Repeat (NoContext)]	Repeat		✓	Rec0003.wav
[Or (NoContext)]	Or		✓	Rec0004.wav
[You can say] Patrick [Ne...	You can say Patrick, Rec...		✓	Rec0005.wav
You can say Patrick [Prev...	You can say Patrick, Pre...		✓	Rec0006.wav
You can say Patrick [Rep...	You can say Patrick, Rep...		✓	Rec0007.wav
[Here is the first]	Here is the first		✓	Rec0008.wav
[Here is the second]	Here is the second		✓	Rec0009.wav
[Here is the third]	Here is the third		✓	Rec0010.wav
[End of sample]	End of sample.		✓	Rec0011.wav
[The following three QA ...]	The following three QA ...		✓	Rec0013.wav

Editr...	Full Transcription	Tag	Wave
You can say	You can say Next Previous or Repeat	BadContext	Rec0000.wav
Next	You can say Next Previous or Repeat	BadContext	Rec0000.wav
Previous	You can say Next Previous or Repeat	BadContext	Rec0000.wav
or	You can say Next Previous or Repeat	BadContext	Rec0000.wav
Repeat	You can say Next Previous or Repeat	BadContext	Rec0000.wav
Next	Next	NoContext	Rec0001.wav
Previous	Previous	NoContext	Rec0002.wav
Repeat	Repeat	NoContext	Rec0003.wav
Or	Or	NoContext	Rec0004.wav
You can say	You can say Patrick Next Patrick Pr...		Rec0005.wav
Next	You can say Patrick Next Patrick Pr...	GoodContextInit	Rec0005.wav
Previous	You can say Patrick Next Patrick Pr...	GoodContextMid	Rec0005.wav



The Speech Debugging Console provides the ability to watch information sent and submitted in a debugging session

The Speech Prompt Editor is an easy way to quickly create, validate, and edit prompts for the developer and voice talent



The Telephony Application Simulator lets developers debug voice-only speech applications on their desktop without requiring Speech Server installation



Telephony and Multimodal Applications

As your business implements a mobile device strategy, Microsoft Speech Server is ready to support you. Multimodal applications, or those that mix speech and visual interactions, are supported in addition to telephony-only applications in MSS as follows:

- Processes SALT-enabled HTML within client browsers
- Add-in for Internet Explorer on Tablet PCs and Desktop PCs speech-enables the Microsoft Internet Explorer client browser and supports local or remote MSS server-based recognition
- Add-in for Pocket IE on Pocket PCs speech-enables the pIE client browser and supports remote MSS server-based recognition over 802.11b networks
- DTMF-only IVR telephony applications supported
- Speech-only IVR telephony applications supported
- Combined DTMF/Speech applications supported

Third-Party Product Support

PBX Integration

The following PBXs are supported by Speech Server via the Intel or Intervoice Telephony Interface Manager (TIM):

- Alcatel 4200, 4400
- Avaya Definity, INDeX
- Deutsche Telekom Octopus E300/800
- Ericsson ACP1000, MD110, BC8-10
- NEC NEAX 2400 IMS Series
- Nortel Meridian 1
- Philips SOPHO iS3000
- Rockwell Spectrum, Transcend
- Siemens HICOM, HiPath

Computer-Telephony Integration

For CTI connectivity, the following vendors, products and interfaces are supported by Speech Server via the Intel or Intervoice Telephony Interface Manager (TIM):

- Aspect
- Intel NetMerge™
- Genesys Labs
- Cisco
- OAI, Passageway

Supported Telephony Boards

Microsoft has partnered with industry leaders Intel and Intervoice to deliver two key components for any Microsoft Speech Server solution – Intel Dialogic boards and the Telephony Interface Manager (TIM).

The Intel TIM supports the following Intel boards:

Type	Ports	Telephony Board
Analog	4	D/41JCT-LS
Analog	16	DM/V160/LP
Digital	24	Contact Microsoft Speech Partner Distributors for the 24-Port Solution
Digital	48	DM/V480A-2T1-PCI
Digital	96	DM/V960A-4T1

The Intervoice TIM supports the following Intel boards:

Type	Ports	Telephony Board
Analog	4	DMV/160/LP
Analog	16	DMV/160/LP
Digital	48	DM/V480A-2T1-PCI
Digital	96	DM/V960A-4T1

Telephony Interface Manager (TIM) Software

Microsoft has partnered with industry leaders Intel and Intervoice to deliver a required component for any Microsoft Speech Server solution — the Telephony Interface Manager (TIM). This third-party component provides the following functionality when used with Microsoft Speech Server:

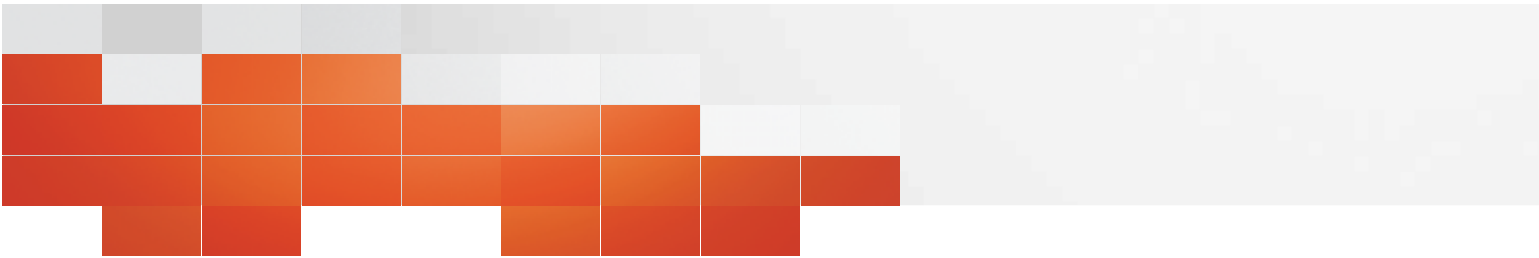
- Controls communication with telephony hardware and the telephony network
- Streams media to and from Speech Engine Services (SES) for recognition and speech output
- Processes call control messaging and call signaling
- Performs DTMF detection

Intel Telephony Interface Manager

The Intel TIM, named the Intel NetMerge Call Manager, isolates application developers and integrators from dealing with the low-level, real-time, driver-level telephony events, messages, protocols, control, and telephony boards. Developers can focus on speech application design and flow because they are shielded from telephony resource management chores. The Intel NetMerge Call Manager works in conjunction with Microsoft Speech Server, providing management and control over Intel Dialogic telephony resources. The call manager enables application development using Microsoft speech tools to control and monitor calls and apply speech resources.

Intervoice Telephony Interface Manager (TIM)

The Intervoice Telephony Interface Manager and Foundation Services enables Microsoft Speech Server to integrate into virtually any call center or telephony environment. It supports Cisco ICM, Genesys, and EADS Telecom PointSpan (OAI) CTI interfaces through the CTI Gateway Web Service, which integrates easily into the Microsoft Visual Studio .NET environment. The Intervoice TIM also provides an integrated Microsoft Management Console (MMC) snap-in that includes support for advanced call routing, outbound hunt groups, dynamic port allocation, enhanced monitoring and reporting, and in-bound call handling.



Microsoft Speech Server 2004 Edition and Feature Matrix

Features	Standard Edition	Evaluation Editions	Enterprise Edition
<24 Telephony Ports	▲	▲	▲
>24 Telephony Ports		▲	▲
Windows Server 2003 Standard	▲	▲	
Windows Server 2003 Enterprise		▲	▲
Single Server Only	▲	▲	
Multiple Server Only, Distributed		▲	▲
For Use in Production Environment	▲		▲
Optimized for 4 and 16 Port Analog Channels and 24 Port Digital Channels	▲	▲	
Optimized for 48 to 96 Port Digital Channels		▲	▲
Supports Third-Party SAPI-Compliant Non-Microsoft Speech Recognition Engines		▲	▲
CTI Vendor Support	▲	▲	▲
PBX Vendor Support	▲	▲	▲
Supports Land Line, Cellular, and GUI-based Devices	▲	▲	▲
Supports Touch Tone (DTMF) and Speech Applications	▲	▲	▲
Supports Telephony and Multimodal Applications	▲	▲	▲
Includes SDK	▲	▲	▲
Supports Intel and Intervoice TIM	▲	▲	▲
License Expires in 180 Days		▲	

System Requirements

To use Microsoft® Speech Server 2004 Standard Edition (for U.S. English only), you need:

- Computer with a 2.5 GHz or faster processor (Microsoft Speech Server 2004 Standard Edition supports up to 4 CPUs on one server)
- Microsoft Windows Server™ 2003 Standard Edition or Enterprise Edition
- 2 GB of RAM or more recommended
- NTFS-formatted local partition with 1.7 GB of available hard-disk space for typical installation; logging requires an additional 20 GB
- Network interface card
- CD-ROM or DVD-ROM drive
- Super VGA (800 x 600) or higher-resolution video adapter and monitor
- Keyboard and Microsoft Mouse or compatible pointing device

To use Administrative Tools, you need:

- Computer with a 600 MHz or faster processor
- Microsoft Windows® 2000 Server with Service Pack 4 or later, Windows 2000 Professional with Service Pack 4 or later, Windows XP Professional with Service Pack 1 or later, or Windows Server 2003 Standard Edition or Enterprise Edition
- 256 MB of RAM or more recommended
- Microsoft Internet Explorer 5.5 or later

To use Telephony Application Services, you need:

- Supported telephony card
- Supported Telephony Interface Manager

To use Microsoft Speech Application Software Development Kit 1.0, you need:

- Computer with a 600 MHz or faster processor
- Microsoft Windows 2000 Server with Service Pack 4 or later, Windows 2000 Professional with Service Pack 4 or later, Windows XP Professional with Service Pack 1 or later, or Windows Server 2003 Standard Edition or Enterprise Edition
- 256 MB of RAM or more recommended
- 500 MB of available hard-disk space
- Microphone
- Microsoft Internet Explorer 6 Service Pack 1 or later
- Microsoft .NET Framework 1.1
- Microsoft Visual Studio® .NET 2003 required for tools only

Actual system requirements will vary based on your deployment configuration, expected load, and the features you choose to install.

To use Microsoft Speech Server 2004 Enterprise Edition (for U.S. English only), you need:

- Computer with a 2.5 GHz or faster processor (Microsoft Speech Server 2004 Enterprise Edition supports up to 8 CPUs on one server)
- Microsoft Windows Server™ 2003 Enterprise Edition
- 4 GB of RAM or more recommended
- Hard-disk space:
 - NTFS-formatted local partition with 1.05 GB of available hard-disk space for Telephony Application Services (TAS) installation; logging requires an additional 20 GB
 - NTFS-formatted local partition with 1.6 GB of available hard-disk space for Speech Engine Services (SES) installation; logging requires an additional 20 GB
- Network interface card
- CD-ROM or DVD-ROM drive
- Super VGA (800 x 600) or higher-resolution video adapter and monitor
- Keyboard and Microsoft Mouse or compatible pointing device

To use Administrative Tools, you need:

- Computer with a 600 MHz or faster processor
- Microsoft Windows 2000 Server with Service Pack 4 or later, Windows 2000 Professional with Service Pack 4 or later, Windows XP Professional with Service Pack 1 or later, or Windows Server 2003 Standard Edition or Enterprise Edition
- 256 MB of RAM or more recommended
- Microsoft Internet Explorer 5.5 or later

To use Telephony Application Services, you need:

- Supported telephony card
- Supported Telephony Interface Manager

To use Microsoft Speech Application Software Development Kit 1.0, you need:

- Computer with a 600 MHz or faster processor
- Microsoft Windows 2000 Server with Service Pack 4 or later, Windows 2000 Professional with Service Pack 4 or later, Windows XP Professional with Service Pack 1 or later, or Windows Server 2003 Standard Edition or Enterprise Edition
- 256 MB of RAM or more recommended
- 500 MB of available hard-disk space
- Microphone
- Microsoft Internet Explorer 6 Service Pack 1 or later
- Microsoft .NET Framework 1.1
- Microsoft Visual Studio .NET 2003 required for tools only

Notes:

Speech Server 2004 Enterprise Edition requires installation of the TAS and SES features on separate, dedicated CPUs. For distributed systems, hardware or software load balancing is required. Actual system requirements will vary based on your deployment configuration, expected load, and the features you choose to install.

www.microsoft.com/speech

© 2004 Microsoft Corporation. All rights reserved. Microsoft, Visual Studio, Windows, the Windows logo, Windows Server, and Windows Server System are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries. The names of actual companies and products mentioned herein may be the trademarks of their respective owners.

Microsoft®