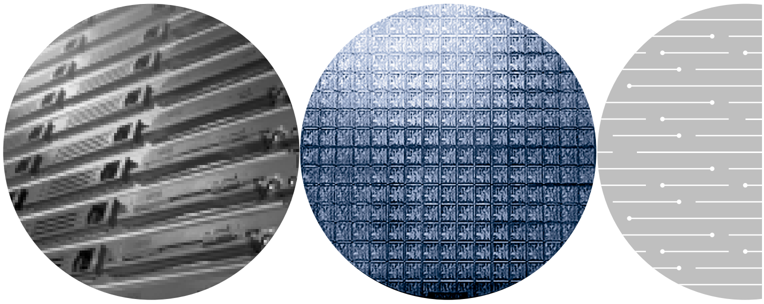
Whitepaper

Intel in
Communications

# Next-Generation Media Processing for the Modular Network

## Using Off-the-Shelf Processors Brings Significantly Lower Cost of Ownership



# Table of Contents

## Executive Summary

This white paper describes a new generation of media processing technology that enables a significantly lower total cost of ownership. Intel calls this technology Intel® NetStructure™ Host Media Processing (HMP) software. HMP software provides basic media services that developers can use to build flexible, scalable, and cost-effective modular network elements for voice and data services.

HMP software moves building media processing solutions from the domain of custom software and specially designed boards to an open software model for standard high-volume computing platforms. This move offers significant advantages.

- **Lower cost of inventory and startup** – smaller initial capital investment
- **Lower development costs** – development systems do not require specialized hardware
- **Lower deployment costs** – software is less expensive to install and configure than hardware
- **Lower sparing costs** – hardware can be used for multiple functions
- **Lower maintenance costs** – maintenance is easier and less training is needed when system configurations are standardized

This paper is primarily technical in nature. It takes a comprehensive view of HMP software technology and its approach to building modular network components.

## Introduction

Over the previous four decades, voice processing has gradually become more computerized, allowing the development of new and more sophisticated services – and lower costs.

The roots of voice band media processing as a technology began in the early Sixties when the electronic switch was first introduced. Previously the telecom network only supported person-to-person interactions between callers and between callers and operators. The electronic switch allowed service to be obtained by interaction with a computer instead. Gradually human-computer interactions became more complex. Touch-tone dialing and automated announcements were introduced, and then fax, modem, conferencing, text-to-speech, voice compression, and speech interaction. These technologies enabled a new set of services that lowered costs (for example, by eliminating human operators) and created new opportunities for service providers and other businesses to add value.

Parallel to the evolution of services, the infrastructure that supported voice media processing evolved as well. The *first generation* saw the introduction of integrated digital signal processing (DSP) silicon across multiple boards and shelves in a proprietary telephony switch. This equipment was very expensive, and new services could take up to five years to implement from concept to deployment because they could only be enabled through either a hardware or proprietary software upgrade, both of which were very costly.

In the 1980s, a *second generation* of infrastructure technology called computer telephony (CT) brought a new open architecture to voice media processing. This technology integrated standard, highly optimized computing boards into open, standardized computing platforms. The cost of service platforms was considerably reduced because the platforms were now based on standard, off-the-shelf computing and communications components. New services were deployed on a multivendor architecture largely driven and differentiated by software rather than single vendor hardware and proprietary software solutions. This dramatically lowered barriers to deploying network services and stimulated a wave of application and service innovation. Businesses created self-service call centers for their customers. Service providers began offering new services to both businesses and consumers, such as calling cards, prepaid phone cards, and voice mail.

Today we are on the verge of a *third generation* of media processing technology made possible by the exponential improvement in general purpose processor performance over many years originally forecast in “Moore’s Law.” In 1965, Gordon Moore of Intel predicted that the number of transistors per integrated chip would double every 18 months. His prediction has proven correct, and Intel believes that Moore’s Law will continue to hold true through the end of the decade, taking into account the current increases in processor density, complexity, and speed.[1]

The newest, most powerful Intel® processors enable a standard computing platform to cost-effectively perform the media processing functions done previously only by

[1] ftp://download.intel.com/labs/eml/download/EML_opportunity.pdf

the special-purpose processors used in the first two generations of this technology. Because of the costs and deployment barriers for media processing will decrease substantially.

Intel's third generation technology for media processing is called Intel® NetStructure™ Host Media Processing (HMP) software. This technology moves algorithms and control software, which previously ran on a proprietary board, to standard computing platforms. Startup costs and total cost of ownership for telecom services will shrink dramatically. Currently application developers need to buy boards and a computing platform to begin developing new telecom applications. With HMP software, developers can start work on new telecom applications after a simple software install. Such a reduced startup cost, along with emerging Internet-telecom integration, promises to spawn a new wave of innovative, cost-effective telecom applications.

## Purpose

For several years, telecom network service providers have been moving to modular open components because of the powerful cost-benefit incentive in using standardized, high-volume hardware and open software in a standard computing model. The ability to save considerably on infrastructure, coupled with the power to create new and innovative services, is a powerful motivation too tempting for the telecom industry to ignore.

A move from media processing on custom, integrated boards to an open software component based on standard, high-volume processors is now feasible. What form is this transition taking? How does HMP software support traditional voice services? What is the future of embedded DSP architecture? How does HMP software enable the deployment of new innovative services? What are the most cost-effective platforms for deploying HMP software? This paper will answer these questions in the following sections.

## Market Segment Opportunities and Challenges

The idea of using HMP software in standard computing systems gained significant visibility with the advent of Voice over Internet Protocol (VoIP) technologies several years ago and the current evolution to a modular next generation network architecture. VoIP encouraged a sizeable group of technologists from the computer and data processing industries to apply their skills to developing solutions for the $250 billion[2] telecom market segment, which was traditionally closed and proprietary. The new services that resulted were more cost effective and flexible, and were based on the concept of a modular telecom architecture with standard protocols and programming interfaces.[3] Industry groups, such as the Internet Engineering Task Force (IETF), created a number of open protocol and programming standards, which further expanded and reinforced support for this architecture. Modular network elements enable basic telecom services, such as local, long distance, and private branch exchange (PBX) switching, to be delivered according to an open computer and data model. Two elements of special interest in relation to the modular network are media gateways and media servers.

Significant opportunity for basic services in the open standard media gateway and media server market segment is very likely in the next few years. In addition, new high value services that are speech-activated and handle Web-based information for businesses, such as electronic assistants and voice portals, further expand the market segment opportunity for modular network elements. The Kelsey Group[4] and Datamonitor[5] expect the end customer opportunity in just these speech-related market segments to grow from under $1 billion today to approximately $4 billion by 2007. Datamonitor also predicts that nearly 80% of these solutions will be built on open platforms by 2007, compared to less than 20% today.

The standard interfaces that have emerged promise to enable the programmers who create eCommerce applications to link telecom services into their current applications and develop unique telecom-only applications. These standards, such as Voice eXtensible Markup Language (VXML), Speech Application Language Tags (SALT), and Standard Interface Protocol (SIP), are defined in the open IETF process, and then linked up with the standardization sector of the International Telecommunication Union (ITU). The goal of the standards is to reduce the deployment time for new converged communications applications from years to months, which aligns with the goals of Information Technology (IT) professionals.

[2] Merrill Lynch, April 2002
[3] "Modular Network Voice Building Blocks" (2002) available at http://www.intel.com/network/csp/pdf/7299.pdf
[4] Kelsey Group, *Voice Ecosystem 2002: Outlook and Planning Guide*, May 2002
[5] Datamonitor, *Voice Portals and Applications*, June 2002

Moving telecom media processing functions into computing silicon and industry-standard computing platforms promises to allow telecom equipment suppliers and enterprises to significantly improve performance. HMP software on computing silicon will be cost effective in low-end deployments that utilize a single computing platform for both application and media processing and also in high-density solutions that require disk storage and support advanced media processing such as speech recognition and text-to-speech (TTS).

Telecom media processing solutions, which focus on very high density streaming, may still be more cost effectively deployed on dedicated DSP silicon. These carrier density functions include

- Media gateways deployed in central offices
- Media servers that process compressed voice streams
- Wireless equipment for baseband signal processing

The importance that these markets segments place on space, power, and heat dissipation can best be met today by silicon designed specifically for these functions. However, continued advances in processing power and improved hardware and software may make specialized silicon obsolete within the next few years.

## Host Media Processing Technology Overview

Intel offers two processor families that address the media processing market segment:

- Processors designed and optimized specifically for digital signal processing
- General-purpose processors enhanced with digital signal processing capabilities

### Specialized Intel® Processors

The processors in the first category above are part of the IXS program, and are specialized DSPs for packet processing. For example, the Intel® IXS1000 media signal processor is a carrier-class, digital system-on-a-chip that employs Voice over Packet (VoP) technology to deliver high-quality voice, fax, and data communications over next-generation optical networks. With the IXS1000 solution, carriers are able to move voice traffic from circuit-switched networks to more efficient, packet-based VoIP and Voice over ATM (VoATM) networks.

For more information on this technology, see the IXS1000 white paper "Optimized High-Density Voice over Packet (VoP) Architecture for Next-Generation Networks" that is available as a resource link from http://www.intel.com/design/network/products/wan/vop/ixs1000.htm.

### General-Purpose Processors

This paper focuses on the second category: general-purpose processors enhanced with DSP capabilities. Intel is developing software components to enable host media processing that uses the signal processing features in the Pentium processor, Itanium architecture, and the Intel® XScale™ microarchitecture. This program will bring the price-performance advantages of general purpose computing platforms to telecom media processing solutions.

The first phase of the program focuses on telecom processing and will begin by using the Intel 32-bit computing platform (IA-32): Celeron®, Pentium, and Intel® Xeon™ processors. Intel's 64-bit architecture (IA-64) on the Itanium processor promises to greatly improve the power and scalability of deployments of HMP software in the future. Itanium processors have an enhanced digital signaling processing instruction set that promises to make it extremely attractive as a platform for high-end telecom media processing.

The Intel XScale microarchitecture, with its low power needs and excellent heat dissipation, is well suited for use in low-end solutions, such as access gateways for the residential and small office market segments delivered by service providers through core network applications.

**History of DSP Capabilities on Off-the-Shelf Processors**

Intel first introduced basic DSP functions for its processors with MMX™ technology on the Pentium processor. MMX (Multi-Media eXtensions) technology began a major enhancement to the Intel architecture that was designed to accelerate multimedia and communications software. This technology essentially embedded basic DSP capabilities in the Pentium processor and included new data types and 57 new instructions to accelerate calculations common in audio, 2D and 3D graphics, video, speech synthesis and recognition, and data communications algorithms. As the number of transistors on a chip has grown, Intel has continued to enhance the DSP ability of the Pentium processor.

The capability that was created with MMX technology has been expanded with Streaming SIMD Extensions (SSE). Streaming SIMD (Single Instruction-Stream Multiple Data-Stream) extensions are instructions that reduce the overall number of instructions required to execute a particular program task. As a result, they can boost performance by accelerating a broad range of applications, including video, speech, and image, photo processing, encryption, financial, engineering, and scientific applications. Intel® NetBurst™ microarchitecture adds 144 new SSE instructions, which are known as

SSE2 and are available on the Intel® Pentium® III and Celeron processors.

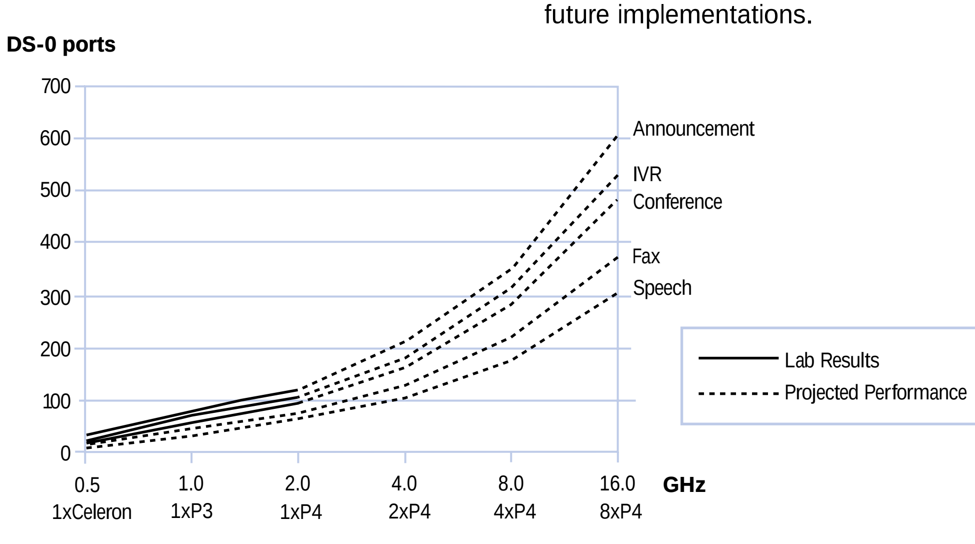
Today, the Pentium and Itanium processors perform floating point operations, matrix operations, fast Fourier transform, finite impulse response, and multithreading. In addition, Intel provides Intel® Integrated Performance Primitives[6] to optimize digital signal processing.

**Preliminary Test Results**

The introduction of DSP capabilities on the Pentium processor has made it an extremely cost-effective technology for deploying host media processing functions. According to preliminary estimates, Intel anticipates that HMP software will enable a single 1GHz Pentium processor to replace two 150 MHz DSPs. From this baseline, Moore's Law improvements to the Pentium processor will enable HMP software at densities of DS-3 and even OC-3. Figure 1 shows projected performance and lab results during which HMP software used 50% or less of processor capacity while an application was running for six major communication media processing functions.

1. **Announcements** – Streaming µ-law or A-law audio data from a file on a hard disk and converting into a telecom audio stream through a linear player resource
2. **Interactive Voice Response (IVR)** – Combining announcements with a dual tone multifrequency (DTMF) or touch-tone signal detector algorithm. This function includes barge-in, a feature that stops the announcement when a DTMF digit is detected or when speech is detected.
3. **Conferencing** – Bridging voice streams in a three-way conference
4. **Transcoding** – Compressing voice, based on the G.729a and G.723.1 compression algorithms
5. **Fax** – Translating from a T.30/T.17 modem stream to fax Tag Image File Format (TIFF) files
6. **Speech** – Performing a range of speech technology functions, including TTS, speech menus, and large vocabularies *(Note: Included only for comparison)*

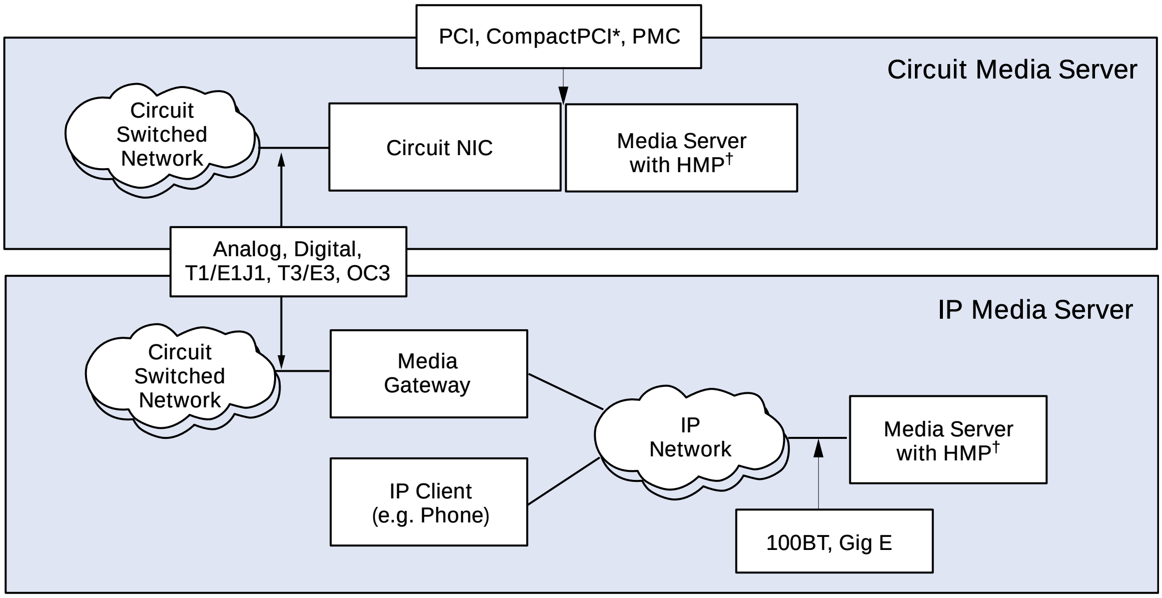
The HMP software program at Intel has implemented, optimized, and tested announcements, IVR, and conferencing algorithms on platforms with a Celeron microprocessor and Pentium processors ranging from 566 MHz to 2 GHz. The solid lines in Figure 1 represent the lab results. Intel is planning an implementation of the algorithms that scales the technology on higher speed Pentium processors and Pentium processors on two-way, four-way, and ultimately eight-way servers, Xeon processors, and Itanium processors. The dotted lines in Figure 1 project Intel's planned results for some of these future implementations.




Celeron=Intel® Celeron® processor
P3=Intel® Pentium® III processor
P4=Intel® Pentium® 4 processor

**Figure 1. Lab Results and Projected Performance**

[6] See http://developer.intel.com/software/products/ipp/ipp30/




† Intel® Netstructure™ Host Media Processing software

**Figure 2. Media Server Deployment Architectures**

## Deployment Environments

The modular next-generation network combines voice and data in a single packet infrastructure. As a result, two main architectures are currently feasible for deploying media services: a direct circuit-based network interface and a packet-based IP network interface. These two architectures are shown in Figure 2. HMP software can provide media processing in both.

Packet-based IP network deployments only require Ethernet network interfaces. New, and more recently traditional, service providers are moving to standard 100BaseT and Gigabit Ethernet as the primary interface for both Internet backbone services and telecom voice connections. Service providers are deploying media gateways to convert the circuit network voice stream to an Ethernet-based infrastructure using VoIP. The ubiquity of Ethernet has necessitated equipping standard Web computing platforms, and now desktop and laptop computers, with Ethernet interfaces. As a result, an Ethernet telecom media server can be deployed on a standard Web computing platform with HMP software alone and without any additional voice processing hardware. These Ethernet networks allow Ethernet-based telecom media servers to provide voice media processing for both legacy circuit clients and IP-based clients.

The same type of Web server platform just described is currently being used for call processing functions in softswitch applications. Telecom media processing services can easily be added to these networks with HMP software running on the same server platform as the softswitch. This new type of deployment would replace the specialized media processing platforms currently in use. The economics of using homogeneous hardware architecture are very compelling. Total cost of ownership is significantly reduced for both network and enterprise service providers because the costs of sparing, training, and the integration of management systems are so much lower. Such a deployment also speeds time-to-market and accelerates the innovation cycle by allowing new features to be deployed through a software upgrade alone.

Deploying a media server with a connection to the traditional circuit switch network or an asynchronous transfer mode (ATM) network requires the addition of a Circuit network interface card (NIC). A Circuit NIC provides either the circuit switched line interfaces (analog [loop start, ground start] or digital [ISDN BRI/BRA[7]]), or the circuit-switched, high-capacity trunk interfaces

[7] Integrated Services Digital Network Basic Rate Interface/Basic Rate Access

(T-1/E-1/J-1, E3/T3, OC3). A typical Circuit NIC includes the following basic functions:

- A line interface unit (LIU) to terminate the copper wires or fiber[8]
- A framer for digital interfaces to terminate the layer 1 protocols or a COder DECoder (CODEC) for analog interfaces for conversion from analog voice signals to digital signals
- A standard Peripheral Component Interface (PCI) streaming interface to the computing platform in PCI AT[9], CompactPCI*, or PCI Mezzanine Card (PMC) form factors

For high capacity trunk interfaces, the Circuit NIC could also include specific components to optimize time division multiplex (TDM) channelization data and perform high-level data link control (HDLC) controller function processing for ISDN and Signaling System 7 (SS7) signaling.

The Circuit NIC streams data to the computing server platform with very low latency (typically less than 8 ms). The computing server platform then performs the media processing and, if applicable, call control signaling processing as well.

In support of the growth of data networking, high density (T-1/E-1/J-1, T3/E3, OC3, ATM) NICs are becoming more common in the market segment. Simple Circuit NICs enable low-cost plug-ins into standard routers as a Wide Area Network (WAN) interface. They also allow computing platforms to perform routing functions. As the design of Circuit NICs becomes standardized, their evolution will be similar to Ethernet NICS in their market segment, but more limited. This evolution should drive down the cost and move Circuit NICs towards the same “plug and play” compatibility that is now available with Ethernet NICs. Intel’s standard “plug and play” interface for processing Circuit NIC data is described in greater detail in the next section.

[8] For example, Intel® LXT350/351 for T1/E1 short-haul transceivers could be used. See *http://www.intel.com/design/network/products/wan/tecarrier/lxt350-351.htm* for details.

[9] Peripheral Component Interface Advanced Technology

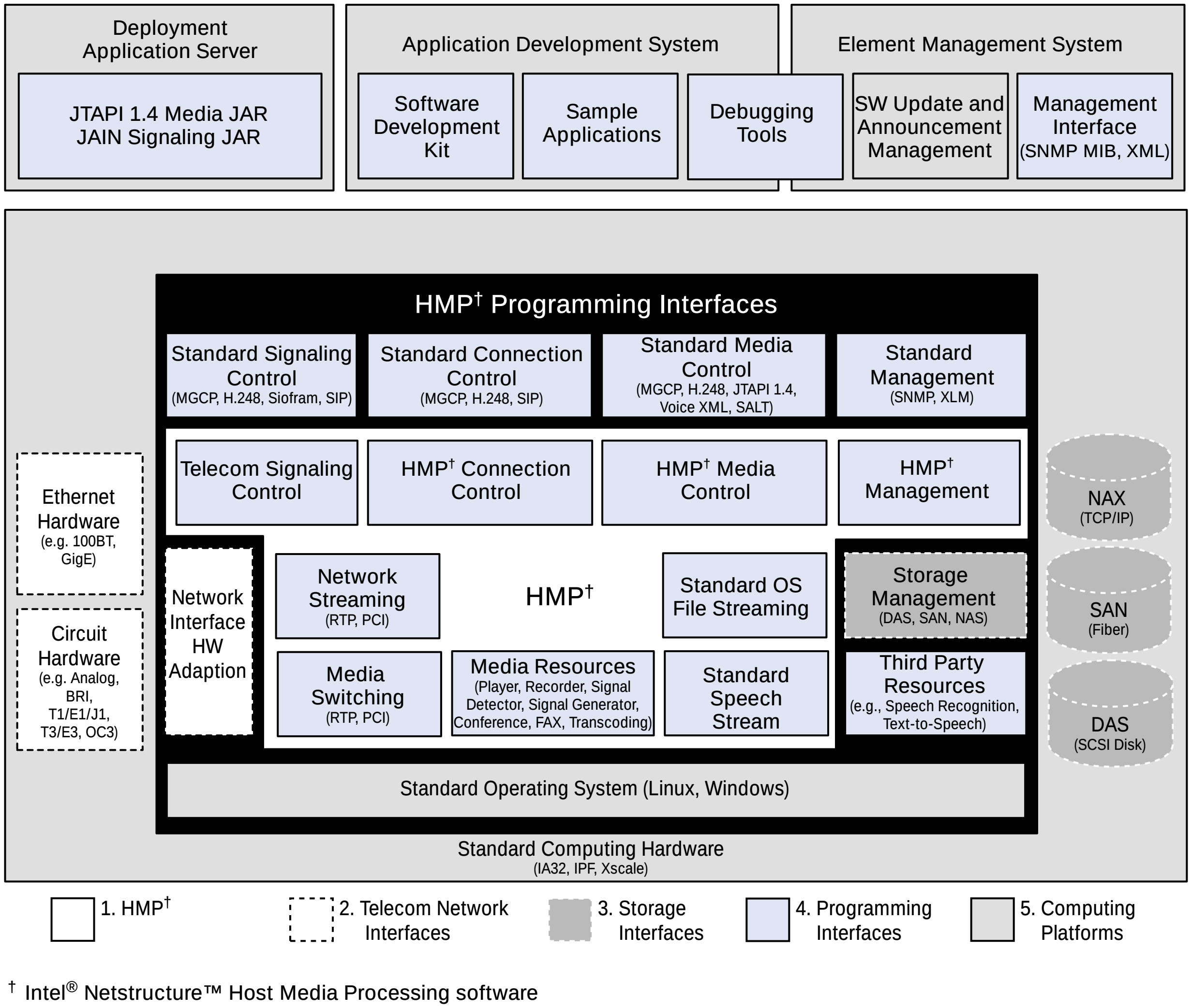



**Figure 3. HMP Software Block Diagram**

## HMP Software Reference Architecture

Intel has developed the HMP software block diagram in Figure 3 as a blueprint for the components in a telecom media server solution. This architecture defines the requirements for the interfaces and functionality to support both circuit-based and packet-based media servers.

The reference architecture has five major functions.

1. HMP software
2. Telecom network interfaces
3. Voice and data storage interfaces
4. HMP software programming interfaces
5. Standard computing platforms

Each of these functions is described in detail in the following sections.

### HMP Software

HMP software supports two major functions: voice media processing algorithms and software voice switching.

#### Voice Media Processing Algorithms

The voice media processing algorithms support the functionality for voice processing resources: announcements, IVR, conferencing, fax, and transcoding. The initial introduction of HMP software is focusing on the algorithms and corresponding resources listed in *Appendix A: Media Capabilities, Resources, and Algorithms.*

However, developing algorithms is insufficient. To create a robust voice media-processing framework, two other areas require attention.

1. Tuning the algorithms for network deployment
2. Optimizing the algorithms for the platform

Intel's algorithms are derived from Intel® Dialogic® and Intel NetStructure products. These algorithms have been tuned for use on telecom networks through millions of ports deployed in more than 100 countries over the last 20 years. Intel is porting this technology to the Celeron, Pentium, Itanium and XScale processor families.

Intel optimizes these core algorithms by recoding the implementations in C to most effectively utilize the MMX and SSE technologies in Pentium and Itanium processors. Optimization can produce up to a tenfold increase in the performance of the algorithm over the initial C code version. For example, by rewriting the basic C algorithm to utilize SSE instructions, Intel expects to increase the efficiency of the G.729a algorithm from 200 MHz to less than 20 MHz per instance of the resource.

Intel is also working with speech technology vendors to reduce costs and simplify the deployment of speech applications. Applications based on speech recognition and TTS technology (e.g., speech IVR, voice portal) provide compelling benefits in terms of operational efficiency and competitive differentiation. An Intel technology called continuous speech processing is an example. This technology optimizes voice activity detection (VAD) algorithms to meet the stringent performance requirements of speech recognition. VAD only streams data to the speech engine when actual human speech is detected, thus optimizing the number of millions of instructions per second (MIPS) that a Pentium processor requires to support speech recognition. According to preliminary estimates, Intel anticipates that a 1 GHz Pentium III processor can support over 100 channels of streaming to a speech engine. The result is an increase in densities and a lower cost when deploying speech-enabled solutions.

**Software Voice Switching**

Intel's optimization efforts extend beyond core media processing algorithms to include the framework's media switching core. The core supports switching voice data to and from resources, to and from files, and to and from third-party resources (e.g., speech recognition and TTS).

Media switching must move data to and from the network interfaces with an extremely low latency. Total end-to-end latency must be less than 200 ms for real-time-sensitive services such as conferencing so that a human user does not perceive a time delay. The HMP software framework will support conferencing with less than 50 ms of latency to ensure an end-to-end latency of below 100 ms.

Another optimization example is VoIP Real-Time Transfer Protocol (RTP) processing on an Ethernet media server. Intel is optimizing its Ethernet driver implementation to reduce the overhead of RTP-based streaming and switching in the platform.

## Telecom Network Interfaces

Network interfaces provide the connection to the circuit network or client device (e.g., analog or digital phone) as defined in *Deployment Environments* (see page 5). The HMP software reference architecture defines two types of network interfaces.

1. Ethernet network interface card (Ethernet NIC)
2. Circuit network interface card (Circuit NIC)

These interfaces stream data to the computing platform through a standard PCI driver that is part of the computing server operating system (OS).

**Ethernet Interface**

The Ethernet NIC streams voice data in and out of the computing platform based on the standard IETF RTP.[10] The media-processing framework unbundles the RTP data and streams the media content to the media processing resources under the control of the application. Data originating from an announcement storage or media resource is likewise streamed from the media-processing framework to the Ethernet NIC. (Most currently available computing servers include dual Ethernet NICs.) As carriers and enterprises move to a homogenous Ethernet architecture for both voice and data, voice-enabled computing servers can plug into the networks with no additional hardware. This converged network will make implementing solutions faster and less expensive.

**Circuit Interface**

The Circuit NIC streams voice data in and out of the computing platform through the PCI bus in buffered TDM form. As with the Ethernet interface, the media-processing framework unbundles the TDM data and passes it to the media-processing framework. Trunk circuit interfaces typically require the interface to perform echo cancellation before engaging a specific resource like IVR, conferencing, messaging, or speech recognition. Intel expects that a 1 GHz Pentium III processor will support over 150 channels of echo cancellation.[11]

[10] Schulzrinne, H., et. al., *RTP: A Transport Protocol for Real-Time Applications*, RFC 1889, January 1996

[11] Assumes 100% utilization of the Pentium processor with no other media processing.

The circuit media interface also streams digital signaling protocols (ISDN or SS7). These protocols require HDLC or Message Transfer Part 1 (MTP1) processing and level 2 Link Access Protocol-D/Message Transfer Part 2 (LAP-D/MTP2) processing fix order/mixing layers. Terminating these protocols is done through one of the following architectures:

- Dedicated hardware on the circuit media interface board for level 1 processing
- Level 1 software processing on the computing platform

Utilizing dedicated hardware on the Circuit NIC has the advantage of offloading the overhead of the level 1 processing from the computing platform. Intel expects to support both architectures. In both cases, the signal protocol is terminated at level 2, and the level 3 packets are sent or received from the application through either the signal control interface or a standard protocol such as Simple Computer Telephony Protocol ISDN User Adaptation layer (SCTP IUA) for ISDN or Simple Computer Telephony MTP3 User Adaptation layer (SCTP MxUA) for SS7.

A standard Circuit NIC interface will allow service providers to choose between multiple Circuit NIC interface types (e.g., analog, BRI, T-1/E-1/J-1, DS-3, etc.) and ideally supply vendors with "plug and play" compatibility with the HMP software architecture.

## Voice and Data Storage Interfaces

Most telecom media processing applications require the ability to play announcements. For example, network call centers for automated 800 services can require storage of over 100,000 announcement files that allow for multiple language support and multiple service support. Files are also the basis of storing voice mail and fax mail in messaging applications. To facilitate the fulfillment of these requirements, HMP software supports the streaming of data from files and the storage of data to files on standard computing platform disks or flash.

HMP software streams data from a file using standard OS file system primitives. This enables the operator to choose the type of storage used.

## HMP Software Programming Interfaces

To allow integration with various industry architectures and mapping across standard industry interfaces, programming interfaces are divided into four areas:

- Call signaling or call control
- Connection control
- Media control
- Management or administration

The programming interfaces are designed to provide two levels of integration across these four functional domains. HMP software provides direct C-language programming interfaces to enable original equipment manufacturers (OEMs) to integrate application components on the computing platform. This allows OEMs to install their applications on the same platform as the algorithms and provides backward compatibly with applications written for the Intel® Dialogic® DM3 architecture.

In addition, industry-standard programming interfaces can be provided that enable "plug and play" interfaces to solutions such as softswitches and application servers. Examples of this and other levels of integration can be found in the *Getting Started with HMP Software* section below. A high level mapping of application control interfaces and detailed media services is outlined in *Appendix B: Programming Interface Capabilities.*

### Call Signaling or Call Control

Call signaling or call control provides the services that interact with the network to set up a voice session between two voice endpoints. Using one of the five major voice network call signaling protocols typically does this: inband (e.g., R1 or R2), ISDN, SS7, H.323, or Session Initiation Protocol (SIP).

Call control is an optional part of HMP software. It is defined in the media server reference architecture to the extent necessary to specify a reference to media processing resources and enable the streaming of the call signaling data associated with the media stream to the application level call processing functions (e.g., softswitch). In addition, it is useful to define call control to support an all-in-one-solution. (See *Getting Started with HMP Software,* page 12, for more information.)

In general, the architecture assumes that an application will interact with call control outside of HMP software and will define the association to set up a voice connection to the media processing resource through connection control. This offers the OEM and service provider optimal flexibility to integrate any major call control architecture as required by an application's specific features.

**Connection Control**

Connection control gives the application control when linking the voice stream (either circuit digital signal level 0 [DS0] or IP RTP session) to a media resource by controlling the HMP-software-switching configuration. The connection control interface also allows the application to create a connection between two voice streams (e.g., two individuals in a call). HMP software provides a connection control interface based on the DM3 architecture model. This model allows granular control when setting up RTP sessions, Circuit NIC connections, and resource connections in both uni- and bi-directional modes.

With advantages very similar to the those of IP, a software-based architecture can move data by passing pointers to shared memory segments. This flexibility simplifies the development of advanced applications for conferencing, call center, and other services by enabling a variety of connection architectures between resources. For example, a configuration of 100 DS-0s could include 100 ports for IVR and only 20 ports for conferencing. The flexible switching in the software solution allows the conferencing ports to be brought in dynamically and trade the MIPs of a Pentium processor with the IVR resources. In a hardware implementation, conferencing is typically dedicated to either all the resources on a board or a percentage of them. Resources cannot be traded between applications easily.

Also, by using software-based switching and an Ethernet fabric, architecture can be scaled for a large switching implementation using standard computing hardware. For example, instead of configuring conferencing as a shared resource with IVR on a server, conferencing can be configured as a separate resource in the network. Ten different IVR servers could connect to the conferencing server over IP when a conferencing resource is required.

**Media Control**

Media control is central to the software providing the programming interface for all the HMP software resources and algorithms defined in *Appendix A: Media Capabilities, Resources, and Algorithms*. The direct media control interfaces are fully compatible with DM3 series application programming interfaces (APIs) for media written in C. Thus, a rich set of applications developed for Intel® Dialogic® products during the past 20 years can now benefit from a software-only architecture when HMP software is used.

The direct media control interfaces provide the granular level of algorithmic control required for highly tuned applications (e.g., engage a specific echo canceller). Included is the control of required algorithmic thresholds and configuration, and input and output (e.g., DTMF digits).

In the area of standard interfaces, five telecom media control-programming interfaces have been defined by the industry: Media Gateway Control Protocol (MGCP) Audio Server, H.248 Annex M (H.248.9), Java Telephony API (JTAPI) 1.4 Media, VXML, and SALT.

MGCP and H.248 are connection control protocols enhanced to support media services that could be deployed on a media gateway or media server. In the case of next-generation softswitches, the enhancements allow the use of one protocol to control all network elements: media gateways or media servers. MGCP and H.248 are focused on the core media functions required by the legacy telecom network (e.g., network announcements, operator services, and three-way calling). MGCP interfaces provide a natural mechanism to explicitly identify coders for transcoding from compressed speech formats to uncompressed formats.

JTAPI 1.4 Media, VXML, and SALT are designed to enable advanced telecom services. JTAPI 1.4 Media (ECTF S.410) is a Java interface specifically defined to support both fundamental media services and advanced speech services. It has not yet gained the same level of acceptance as VXML.

VXML and SALT have been defined specifically to support the integration of speech recognition capabilities with Web-based application environments. VXML has gained a large following in the Web services programming model for voice portal and ultimately speech-driven call center applications. SALT extends VXML concepts to include support for standard Web scripting languages (e.g., Javascript) and multimode applications (e.g., an Internet-enabled phone).

The modular architecture and powerful yet highly usable direct programming interfaces of HMP software enable easy integration with any of the standard programming interfaces described above.

**Management or Administration**

The comprehensive and standard management frameworks provided by computing platforms as part of their operating systems provide significant interoperability with both carrier and enterprise management systems. Platforms typically support both Simple Network Management Protocol (SNMP) and eXtensible Markup Language (XML) management for processor performance, and Ethernet interfaces and disks. In addition, Intel® computing platforms also support the IP management interface (IPMI) for monitoring the platform and providing the robust management interface required for lights-out server operation.

HMP software provides support for three critical management areas necessary for reliable operation and effective engineering: performance, fault, and configuration. The management information base (MIB) for telecom network interfaces, storage, and computing hardware are based on the standard industry capabilities included as part of those components.

Intel is also implementing standard MIB bindings for both the Linux* and Windows* operating systems. Intel is focusing on an SNMP binding for performance and fault management and XML over Simple Object Access Protocol (SOAP) for configuration management. In addition, to support management over the Internet, SNMP Version 3 is being targeted as critical to enhanced security capabilities.

## Standard Computing Platforms

HMP software is designed to enable cost-effective utilization of standard high-volume computing platforms. Because of its architecture, the software can scale from low-end small-office deployments to high-end carrier deployments depending on the enabling computing platform on which it is installed. Intel has computing platform reference designs that range from low-cost appliances based on Celeron processors to high-performance four-way platforms with Xeon or Itanium processors. Server platforms are deployed in five primary models:

1. **Appliance** – Small board (227 mm × 209 mm) architectures for a single function
2. **Entry Web servers** – One- and two-way 1U platforms for cost-effective Web server deployments
3. **Telecom NEBs servers** – One- and two-way 1U and 2U processor platforms for telecom deployments requiring NEBs certification
4. **Blades** – One-way PICMG[12] 2.16 6U blades for industrial and telecom applications. These blades will soon be available in the Advanced Telecommunications Computing Architecture (ATCA PICMG 3.x), and will support configurations beyond two-way.
5. **High-end database servers** – Four- and eight-way platforms with Xeon or Itanium processors for performance-intensive applications

Table 1 provides examples of the performance expected with HMP software on the various computing platforms described above. Estimated availability for these deployments is the first half of 2003.

| Platform | Processor | IVR Performance | Conferencing Performance | Web Site Links | Notes |
|---|---|---|---|---|---|
| One-way appliance | Pentium III, 850 MHz | Tested with 32 channels with 45% CPU utilization | Tested with 32 channels with 45% CPU utilization | http://developer.intel.com/platforms/applied/eiacomm/papers/appliances.htm | Intel is planning to introduce a similar reference design supporting circuit interfaces (analog, ISDN, Digital PBX, T1/E1) |
| Two-way server | Xeon, 2.4 GHz | Projected 240 channels at 50% CPU utilization | Projected 220 channels at 50% CPU utilization | http://developer.intel.com/eBusiness/products/server/processor/ | Circuit media server support available in future with PCI NICs (T1/E1/J1) |
| Two-way NEBs server | Pentium III, 1.26 GHz | Projected 128 channels at 50% CPU utilization | Projected 120 channels at 50% CPU utilization | http://www.intel.com/design/cgservers/telecom/index.htm | Circuit media server support available in future with PCI NICs (T1/E1/J1) |
| 2.16 NEBs blade | Pentium III, 1.2 GHz | Tested 64 channels at 35% CPU utilization | Tested 64 channels at 35% CPU utilization | http://www.intel.com/design/network/products/cbp/linecard.htm | Standard PICMB 2.16 blade with Dual GE. 2x versions planned |

**Table 1. HMP Software Test Results and Projected Performance**

[12] PIC Industrial Computer Manufacturers Group

The projected performance numbers quoted in Table 1 were based on the performance data shown in Figure 1.

The broad range of available platforms compatible with HMP software can help meet the challenges identified in *Market Segment Opportunities and Challenges* (see page 2), especially in terms of implementation speed and overall infrastructure cost reduction.

## Getting Started with HMP Software

This section outlines four typical application and deployment scenarios for HMP software.

- Small office PBX
- Carrier IP media server
- Unified messaging or conferencing server

For each of the applications, required elements are identified along with the interfaces between the application elements and the software components. In addition, the configurations present an outline for projected DS-0 resource configuration and density.

The application configurations presented here should be considered a starting point for different combinations of Intel® hardware and software building blocks, network interfaces, and storage devices.

### Small to Medium Office PBX Applications

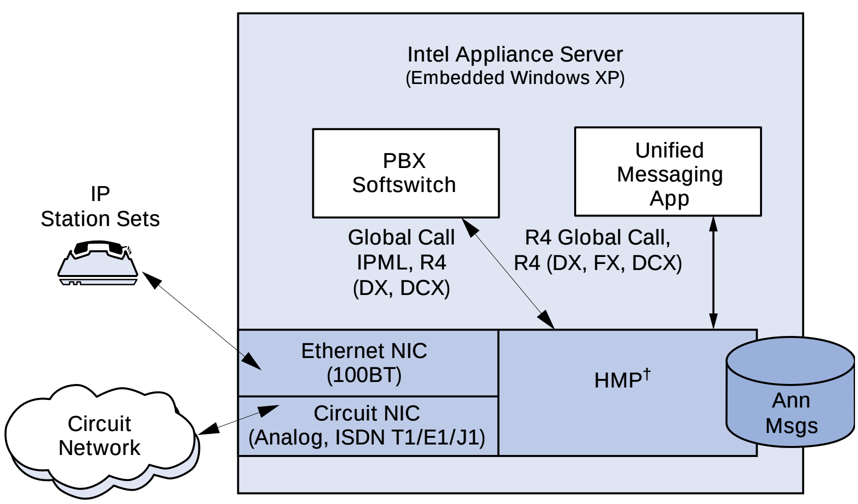
Figure 4 illustrates a cost-effective deployment architecture for small and medium office PBX applications. Two types of applications are typical in this configuration.

1. PBX call control
2. Voice messaging

Both of these applications rely on HMP software for controlling the voice processing with the station sets and circuit network.

A Circuit NIC is required to support either an analog or ISDN connection to the network. The application controls the software through core software programming interfaces. Announcements and messages are stored on a disk on the same server as the software.

The PBX application can be deployed on a low-cost Intel® appliance or on an off-the-shelf desktop computer with an Intel processor. The end user can decide which platform to use based on the value added by the OEM.




† Intel® Netstructure™ Host Media Processing software

**Figure 4. Configuration for Small to Medium Office PBX Applications**

## Carrier IP Media Server

A carrier IP media server configuration requires high performance and reliability to provide core local (class 5), long distance (class 4), and business Centrex* services.

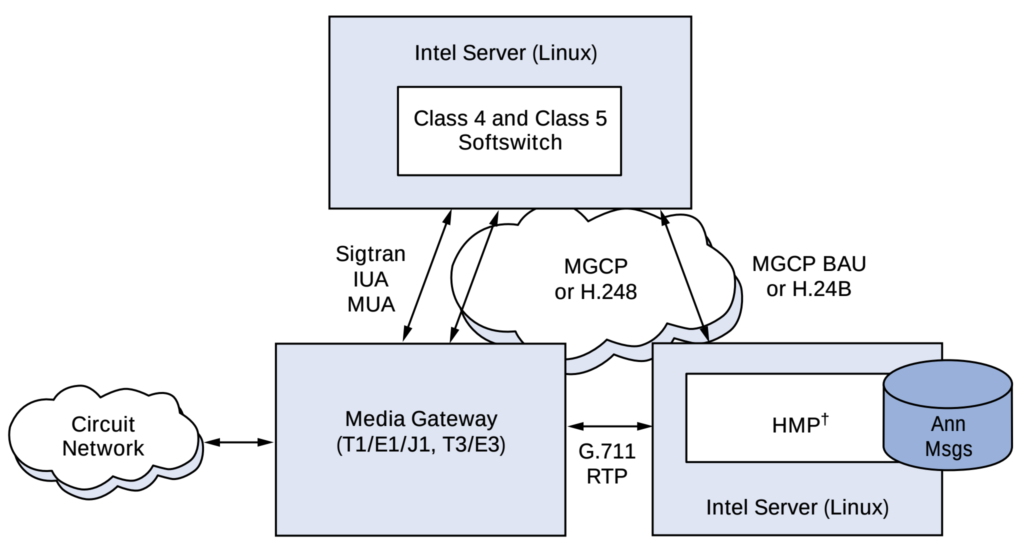
The modular IP architecture for deploying these services, shown in Figure 5, is built on next-generation protocols and splits network services into four elements.

1. A packet backbone, based on IP or ATM
2. A software-based softswitch that implements basic call control and billing functions
3. A media gateway providing the bridge between the circuit and packet networks
4. An IP media server providing basic announcement, IVR, and conferencing functions

The architecture in Figure 5 uses MGCP or H.248 as the controlling interface to the media gateway and media server elements.

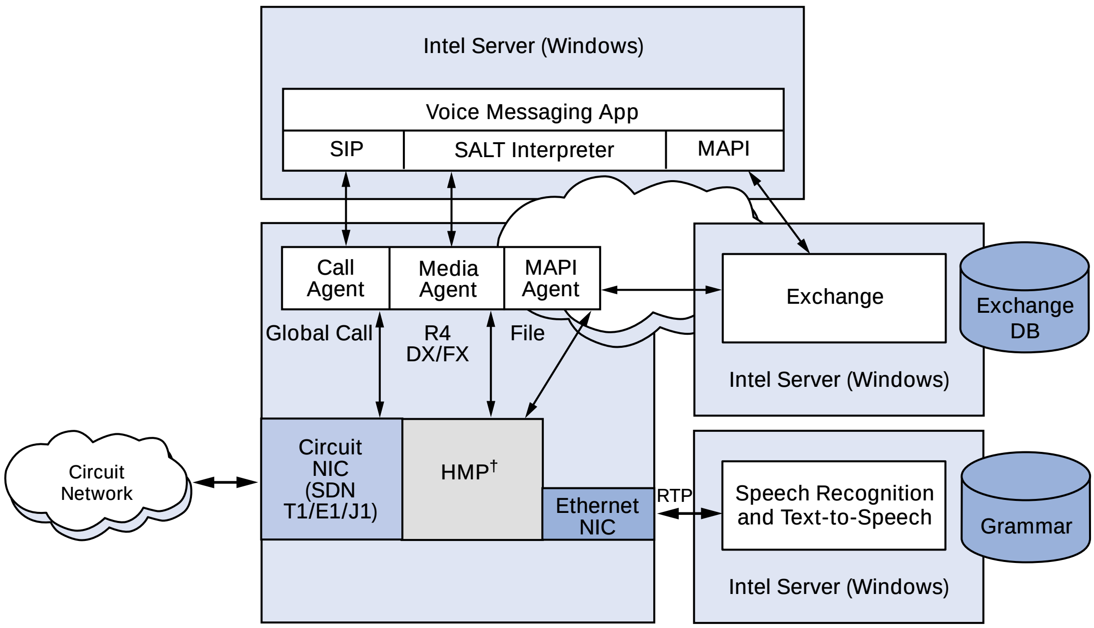
In Figure 5, the softswitch uses HMP software for playing announcements, interacting with touch-tones (DTMF digits), and building conference calls. The softswitch uses the MGCP Basic Audio Package to interact with the media framework, which streams announcements from the local disk through an Ethernet NIC to the media gateway. The software detects incoming DTMF digits from the media gateway RTP stream and provides the ability to mix conference RTP streams from separate and geographically distributed gateways.

According to the projections for the two-way Network Equipment Building Standards (NEBS) server configuration in Table 1, the softswitch can add media services in up to 250 DS-0 field replaceable units (FRU). This allows the service provider to scale media services by adding licenses or platforms.




† Intel® Netstructure™ Host Media Processing software

**Figure 5: Configuration for a Carrier IP Media Server**




† Intel® Netstructure™ Host Media Processing software

**Figure 6: Configuration for Unified Messaging or Conferencing Server**

### Unified Messaging or Conferencing Server

Figure 6 illustrates a configuration for building a unified messaging or conferencing application using HMP software and the Microsoft* .net* architecture.

In .net architecture, Microsoft provides XML programming interfaces for all necessary application interfaces: call control, media control, and message storage. Message storage is based on Microsoft Exchange*, and speech recognition is enabled by a third-party technology linking to the Microsoft platform through the Microsoft Speech Application Programming Interface (SAPI) or directly to the SALT interpreter. The application can be developed completely within the domain of the Microsoft Visual Studio* .net tools.

HMP software provides a complete environment for the telecom media processing functions of the SALT specification, specifically the DTMF, announcement, and conferencing functions. The software also provides a high-performance front-end interface to connect the Circuit or Ethernet NIC to the speech recognition and TTS application servers through the standard RTP streaming interface. Voice messages stream through a Message Application Programming Interface (MAPI) file filter between the HMP software and the Exchange database.

In this configuration, the operator can utilize HMP software to lower the cost of installing and deploying the high-density IVR resources and the small amount of fax and TTS resources that unified messaging typically requires. The platform can also be enhanced with conferencing as a call completion service to follow up on a message. The ability to deploy media services as a software option allows both the efficient use of standard servers and easy platform enhancement as usage and business needs change. HMP software maximizes flexible feature deployment while optimizing the implementation of speech to reduce costs.

## Conclusions

This paper introduces a new technology called Intel NetStructure Host Media Processing (HMP) software. Intel continues to provide new and innovative architectures for deploying new services based on high-performance silicon technologies and Intel architecture.

HMP software implements a deployment, which traditionally uses specialized, low-volume hardware, on off-the-shelf, high-volume building blocks. This modular next-generation technology is expected to dramatically reduce costs for deploying traditional and enhanced telecom services. It should also accelerate the convergence of voice and data services.

## Appendix A: Media Capabilities, Resources, and Algorithms

Table 2 summarizes HMP software capabilities, resources, and algorithms in a media server.

| Switching | Player | Recorder | Signal Detector | Signal Generator | Conference | Fax | Transcoding | Speech Recognition Front-End |
|---|---|---|---|---|---|---|---|---|
| Echo canceller 16 or 32 ms tail (circuit only) | Echo canceller 16 ms or 32 ms tail (circuit only) | | DTMF detector | DTMF generator | Echo canceller 16 ms or 32 ms tail (circuit only) | T.30/V.17 to/from a TIFF file | G.723.1 (5.3k, 6.3k) to/from G.711 | Voice activity detect |
| RTP to RTP | µ-law, A-law, (48k, 64k) | µ-law, A-law, linear | RFC 2833 (IP only) | RFC 2833 (IP only) | n-way summer with VAD | T.38 to/from a TIFF file (IP only) | G.729ab (8k) from/to G.711 | Silence compression |
| RTP to PCI | ADPCM OKI (24k, 32k) G.726 (24k, 32k) file | ADPCM OKI file | Busy, audible, reorder, SIT | Dial tone, busy, ring back, reorder, SIT | DTMF clamping | | G.722.1 (8k, 24k, 32k) to/from G.711 | |
| PCI to PCI | Linear (8/16 bit, 8k) (8/16 bit, 11k) | G.726 file | Modem detect (V.34/V.90) | Call waiting, CNG, answering machine, | Whisper mode | | G.726 (16k, 32k) to/from G.711 | |
| | WAV file, TrueSpeech (8.5k), Realaudio | WAV file | CNG/fax detect | V.23 FSK | Active talker | | GSM-EFR(12.2k), GSM-TIPHON(13k), GSM-MS(13k) to/from G.711 | |
| | Third-party text-to-speech | Third-party text-to-speech | Voice activity detect | | Volume control | | Silence suppression | |
| | Volume up/down | Automatic gain control | | | | | Comfort noise generation | |
| | Speed up/down, skip ahead/ back | Silence compression | | | | | | |

**Table 2: HMP Software Capabilities, Resources, and Algorithms**

The above resources are typically deployed in the following service configurations:

- IVR – Contains a player and a signal detector
- Conferencing – Adds a conference resource to IVR
- Voice messaging – Adds a recorder to IVR
- Unified messaging – Add fax to voice messaging
- Speech IVR – Adds a speech front end and speech engine to IVR

All service configurations require transcoding if the IP RTP stream is not G.711. All circuit streams are µ-law or A-law.

## Appendix B: Programming Interface Capabilities

Table 3 defines component and standard program interfaces that map to HMP software capabilities.

| | Description | Signaling/Call Control | Connection Control | Media Control | Target Application |
|---|---|---|---|---|---|
| **Core framework** | Direct C language libraries for telecom voice media processing control developed by Intel | Call setup and teardown via the Intel® framework interface and features, or the DM3 Global Call interface and features | Switching via DM3 IPML and SCbus interface and features | Player, recorder, signal detector detector and generator, conferencing and fax via DM3 R4 DX, DCB, and FX interfaces and features | OEM implements standard programming interfaces; compatibility with existing applications Intel Dialogic and Intel NetStructure boards |
| **H.323 (H.225 and H.245** | First generation VoIP call control and media control | Call setup and teardown via H.225 Q.931 messages. | Simulated through H.225 messages | DTMF signal detector via H.245 digit messages | Existing H.323 deployments |
| **MGCP** | IETF informational RFC2705bis for controlling VoIP gateways | Call associated signaling control via packages:<br>▪ Trunk<br>▪ Line<br>▪ Signal list | Switching control via packages:<br>▪ Trunk<br>▪ Line<br>▪ RTP | Player, recorder, tone generator, tone detector, conferencing, and fax control via packages:<br>▪ Generic media<br>▪ DTMF<br>▪ Supplementary services<br>▪ Digit map extension<br>▪ Media format<br>▪ Resource reservation<br>▪ Announcement<br>▪ Server script | Class 4, Class 5, and PBX services |
| **H.248** | ITUT protocol for telecom audio and video device control over IP networks | Call associated signaling control via packages:<br>▪ Analog line supervision<br>▪ Network<br>▪ Basic continuity<br>▪ TDM circuit | Switching control via packages:<br>▪ Generic<br>▪ Base root<br>▪ RTP | Player, recorder, tone generator, tone detector, conferencing, and fax control via packages:<br>▪ Tone generator<br>▪ Tone detection<br>▪ Basic DTMF generator<br>▪ DTMF detection<br>▪ Call progress tone generator<br>▪ Call progress tone detection<br>▪ Dynamic tone definition<br>▪ Generic announcement<br>▪ Advanced announcement server<br>▪ Facsimile<br>▪ Text conversation<br>▪ Call discrimination | Class 4, Class 5, and PBX services |
| **Sigtran** | IETF protocol for telecom signaling over IP networks | SCTP IUA, SCTP MxUA | N/A | N/A | Class 4, Class 5, and PBX services |
| **SIP** | IETF protocol for call control over IP networks | Call setup and teardown via SIP user agent messages | Switching simulated by SIP back to back UA architecture | Speech recognition and text-to-speech via IETF CATS; DTMF signal detector via RFC2833 for DTMF | Advanced services such as messaging and call center |

| | Description | Signaling/Call Control | Connection Control | Media Control | Target Application |
|---|---|---|---|---|---|
| **JAIN JTAPI 1.4 Media** | JCP protocol for Java-based telecom media control | JAIN, JTAPI | N/A | Player, recorder, tone generator, tone detector, conferencing, fax, speech recognition | Advanced services and Web application servers |
| **VoiceXML** | IETF W3C protocol for telecom audio and speech control over HTTP sessions | First party – call control XML TBD | N/A | Player (no speed/skip), recorder, signal detector (no call progress), text-to-speech, speech recognition | Web IVR and speech IVR services |
| **SALT** | Industry-proposed protocol audio, speech, and call control over HTTP sessions | Call control using XML and SIP | N/A | Player, recorder, signal detector (no call progress), signal generator, conferencing, text-to-speech, speech recognition | Web IVR, conferencing and speech IVR services; multimode services |

**Table 3: Programming Interface Capabilities**

To learn more, visit our site on the World Wide Web at **http://www.intel.com**.

1515 Route Ten
Parsippany, NJ 07054
Phone: 1-973-993-3000