Volume 10 Issue 01 Published, February 15, 2006 ISSN 1535-864X DOI: 10.1535/itj.1001

Intel® Technology Journal

Converged Communications

The unification of voice, data and video services into converged communications is inevitable. This issue of Intel Technology Journal (Volume 10, Issue 1) examines the architectures, technologies, standards, opportunities, and issues associated with convergence, from service provider networks to the digital home.

Inside you'll find the following articles:

More information, including current and past issues of Intel Technology Journal, can be found at:
http://developer.intel.com/technology/itj/index.htm

Volume 10 Issue 01 Published, February 15, 2006 ISSN 1535-864X DOI: 10.1535/itj.1001

Intel® Technology Journal

Converged Communications

Articles

THIS PAGE INTENTIONALLY LEFT BLANK

Preface

Voice, Video, and Data Services—the Merging of Triplets
By Lin Chao
Publisher, *Intel Technology Journal*

With the wide-spread adoption of Internet Protocol (IP)-based technologies, it is practical and compelling to run voice, video, and data over a single physical data infrastructure rather than over separate networks. This convergence is enabling new enterprise business solutions—converged networks which provide performance to voice, video, and data communications across the enterprise. The eight papers in this issue of Intel Technology Journal (Vol 10, Issue 1) on “Converged Communications” explain why and how the unification of voice, video, and data service infrastructures for both stationary and mobile devices is inevitable. Included are papers that look at actual test beds within Intel’s enterprise that test solutions designed to fit into a complex business setting. These papers describe in detail the compelling usage models which are the key drivers for the unification of voice, video, and data services.

The introductory paper discusses a converged voice, video, and data network architecture and its uses in today’s enterprise. Most enterprises today support at least three separate networks (LAN, WLAN, and voice) for static data, mobile data, and voice. This paper looks at the unification of voice, video, and data service infrastructures for both stationary and mobile devices. Currently there is rapid adoption and migration by telephone, cable, and media vendors and industry to move to IP in order to take advantage of converged networks.

The second paper looks at Session Initiation Protocol (SIP), a signaling protocol which is a widely adopted standard in the industry. SIP enables data and voice convergence for voice, video, Instant Messaging (IM), and other media, and facilitates presence- and location-based services. In this paper we focus on the emerging capabilities of SIP within real-time communication technologies while addressing challenges within the areas of interoperability, security, and enterprise network integration. We describe how the seamless integration of presence-based SIP, VoIP, Mobile IP, SIP mobility support, unified communications, and applications can lead to converged communications.

The third paper examines Advanced Telecom Computing Architecture* (AdvancedTCA*), a set of industry-standard specifications for the next-generation carrier-grade communications equipment. AdvancedTCA incorporates blade (board) and chassis (shelf) form factor optimized for carrier-grade telecommunication applications, with support for carrier-grade features such as NEBS, ETSI, and 99.999% availability. This paper explains the current state of AdvancedTCA and how to mature to the next level as AdvancedTCA industry brings together the component suppliers, system vendors, and service providers in a concerted effort to specify, develop, certify, and deploy interoperable modular solutions.

The next two papers look at test beds around Intel on converging voice, video, and data in Intel’s own enterprise with case studies on some real-world applications. The fourth paper examines a case study based on Intel’s own experience of deploying VoIP with voice quality within a campus and

* Other names and brands may be claimed as the property of others.

converging voice and data on the existing local area network (LAN). The methodology used to enhance end-to-end voice quality includes ensuring bandwidth, enabling QoS, and optimizing IP telephony endpoints. This case study demonstrates that even a basic voice quality and QoS strategy can produce high-quality results with minimal infrastructure upgrades. This project implemented a very basic voice quality and QoS strategy, proving that VoIP can be layered on top of an existing infrastructure with minor upgrades.

The fifth paper explores Intel's recently completed trial of standards-based real-time multimedia collaboration tools running on laptop computers over an Internet Protocol (IP) network. Key goals for the trial included validation of usage models and user benefits while utilizing the multimedia collaboration tools in a production work environment. Trial participants were equipped with a multimedia "softphone" application, a headset, and a Webcam that enabled them to establish high-quality small-group (multiparty) voice and video calls.

The sixth paper describes the role of the media server in traditional Public Switch Telephone Network (PSTN), Internet Protocol (IP), and next-generation networks. We examine the Intel® building-block technologies and their use in developing powerful, cost-effective multimedia communication solutions. We show how processors based on Intel® architecture and Intel® Integrated Performance Primitives (Intel® IPP) can offer world-class performance of media processing algorithms such as audio and video codecs.

The seventh paper summarizes the results of Intel's recent market analysis on unmet consumer communication needs and shows how the collected data suggests the need to develop a general, standards-based framework for digital home communications as well as two new specific device classes. We first show the results of a recent study on unmet consumer needs for communications. We then describe two new digital communication device classes suggested by the CMR data: a Home Communications Server (HCS) and a Digital Communications Adaptor (DCA). The requirements for the proposed digital communications framework are compared with the existing Digital Living Network Alliance (DLNA) framework for digital home entertainment. We conclude with an overview of similar architectural components that will be needed to establish the digital home communications framework.

In the eighth paper, we present Intel's Seamless Collaboration Architecture for VoIP on WLANs using notebooks based on Intel® Centrino® mobile technology. We describe our QoS architecture for VoIP that enables "softphone" applications to take advantage of QoS features over WLANs. We also describe other hardware/software such as array microphones, Intel IPP, and the Bluetooth* wireless coexistence solution.

These eight papers in this issue of Intel Technology Journal (Vol 10, Issue 1) describe the many compelling usage models which are the key drivers for the inevitable unification of voice, data, and video services into truly converged communications.

Foreword

Convergence Now

By Anthony Neal-Graves
General Manager, Modular Communications Platform Division
Digital Enterprise Group
Intel Corporation

Over a decade ago, the voice communications industry started talking about network convergence and its benefits. Converged networks promised cost efficiencies by enabling a single network to support voice, video, and data. More importantly, converged networks promised advanced services and solutions that could change the way businesses and consumers communicate. At that time, the International Telecommunication Union (ITU) was developing the H.323 specification for multimedia communications, and the Internet Engineering Task Force (IETF) was drafting the Session Initiation Protocol (SIP) specification. The world-wide web and e-Commerce were in their infancy, and Wi-Fi and 3G wireless networks did not exist.

How much has changed in a decade! Today, it is impossible to talk about communications without mentioning convergence. A culture of "openness" has permeated the industry. This represents a significant change from the traditional communications industry which was dominated by a small number of companies selling expensive solutions that were difficult to manage, maintain, and upgrade. The Internet and the world-wide web have demonstrated the value of open networks and open systems. They have changed the way we think about communications. There is a groundswell of support for standards in a variety of convergence-related areas including call control, web services, media processing, and network resource control. And let's not forget about the open-source movement, which has expanded beyond operating systems to include important tools, frameworks, and applications.

This open, standards-based network is complemented by ongoing technology advancements which make it easier and less expensive to deploy converged solutions. Multimedia and communications processing tend to be very compute-intensive processes and have traditionally relied on expensive, purpose-build processors. This has kept costs high and stifled innovation for client and infrastructure solutions. Today, we can use general-purpose processors to handle many of these tasks. By using off-the-shelf, modular components for converged solutions, vendors can deliver solutions to the market faster and at lower cost.

The result? New applications supporting Voice over IP (VoIP) and real-time collaboration are being deployed rapidly, and new client platforms are available to support these applications in wireless environments. Application infrastructure is converging, as many providers of web frameworks and tools now support interactive communications technology for instant messaging, voice, video, and collaboration. The service-provider community has aligned on a single architecture for implementing next-generation services in IP-based networks—the IP Multimedia Subsystem (IMS). Convergence is real and this is a great time to be in the communications business.

This issue of ITJ examines the architectures, technologies, standards, opportunities, and issues associated with convergence, from service provider networks to the digital home. We hope you enjoy reading about this exciting area of communications.

Technical Reviewers

Sanjay Bakshi, Mobility Group
Omri Barkay, Information Technology
Monte Bates, Information Technology
Blaine D. Bauer, Information Technology
Chris D. Brinkmeyer, Information Technology
Michael Deisher, Corporate Technology Group
Duncan Glendinning, Mobility Group
Venkat Gokulrangan, Mobility Group
Jim Kennedy, Digital Enterprise Group
Barbara Klein, Sales and Marketing Group
Aniruddha Kundu, Digital Enterprise Group
Charlie Lasswell, Software and Solutions Group
Don Meyers, Software and Solutions Group
Marc Rossi, Digital Enterprise Group
Sanjay Rungta, Information Technology
Glen Shires, Digital Enterprise Group
Shelby Siegel, Information Technology
Michael Stanford, Digital Enterprise Group
Tim Verrall, Information Technology
Gary Veum, Information Technology
Bapi Vinnakota, Digital Enterprise Group

THIS PAGE INTENTIONALLY LEFT BLANK

Enterprise Converged Network–One Network for Voice, Video, Data, and Wireless

Sanjay Rungta, Information Technology, Intel Corporation
Omer Ben-Shalom, Information Technology, Intel Corporation

Index words: Voice over IP, local area network, wireless local area network, asynchronous transfer mode, wide area network, access point, Quality of Service, RADIUS, 802.1x

ABSTRACT

The different components of the trio commonly called "triple play" today (Voice, Video, and Data) were originally developed in different domains, and the networks carrying them were designed and engineered specifically for their requirements. The implication was that different network environments had to be supported concurrently to allow all three services to exist. Asynchronous Transfer Mode (ATM) was the first network technology specifically created to allow for the convergence of data, video, and voice over Wide Area Networks (WANs), but it has failed to be accepted in the Local Area Network (LAN) space due to unproven costs and complexity. Ethernet has come out as the clear dominant LAN technology. With the rapid emergence of mobile networking using Ethernet-based wireless LAN (802.11) technologies the market is exhibiting renewed enthusiasm for communication convergence based on LAN technologies. Recent advancements in security and Voice over IP (VoIP) reliability and quality along with the seamless integration of new WLANs and traditional LANs have provided the technical and business impetus to converge data, video, and voice networks into a single cohesive network service infrastructure. However, supporting varying classes of services and capabilities on LAN and WLAN environments has proved to be very challenging due to strict requirements on IP packet loss, packet delay, and delay variation (jitter). To make convergence of services realistic we are looking at recent advancements in Quality of Service (QoS) algorithms particularly in the areas of process and packet prioritization and scheduling as the main enabler for allowing network architects to overlay voice, data, and video on a shared data network. Furthermore, WLANs (802.11) have become a mainstream capability suitable for the enterprise as they provide converged services while being "always connected." This concept allows the LAN to become an integrated method of connectivity not just for traditional devices such as desktops but also for a large group of mobile computing devices of varying form factors and mobile telephony users presenting a significant and appealing value to businesses. We propose a campus-level LAN in which the three previously separate networks are "converged" seamlessly into one mobility-enabled enterprise network architecture. We estimate that the simplicity of the converged architecture will contribute significantly to the total cost of ownership (TCO) of managing the capabilities independently in a campus for IT.

In this paper we present the enterprise converged network architecture and its uses. We describe the case studies that will be used by ISTG at Folsom for LAN and voice convergence and the plan for a wireless network to integrate with LAN to make it another access media to support all LAN services.

INTRODUCTION

Most enterprises today support at least three separate networks (LAN, WLAN, and voice) to support static and mobile data and voice (Figure 1). Each network was optimally designed to meet different requirements but with the convergence of services the cost and support burden associated with upholding three separate networks are becoming prohibitive. The unification of voice, video, and data service infrastructures for both stationary and mobile devices is inevitable and already termed "quadruple play" by the industry. Specifically the equipment to provide voice services is going to migrate from today's vertically oriented PBXs to IP-based telephony servers supporting call, registration, and enhanced calling services. It is anticipated that core telephony services will integrate and blend with video and text-oriented layered services where network service infrastructure convergence is feasible from an operational and cost perspective. Since voice infrastructure will ride over IP, its quality will depend on QoS-enabled IP core network services for delivery of real-time transport. User and operational advantages are anticipated through a more flexible system based on a PC or server-based model for rapid evolution via voice and data service integration.

The development and integration of rich and integrated applications over a common set of layered services will also enable the unification of electronic mail, voice and text messaging (e.g., IM) over common users' devices and a converged network infrastructure. Nevertheless, matching the scalability, cost, and reliability advantages which the enterprise maintains with its current environment, will be the single biggest challenge in realizing a voice and data convergent communication vision.

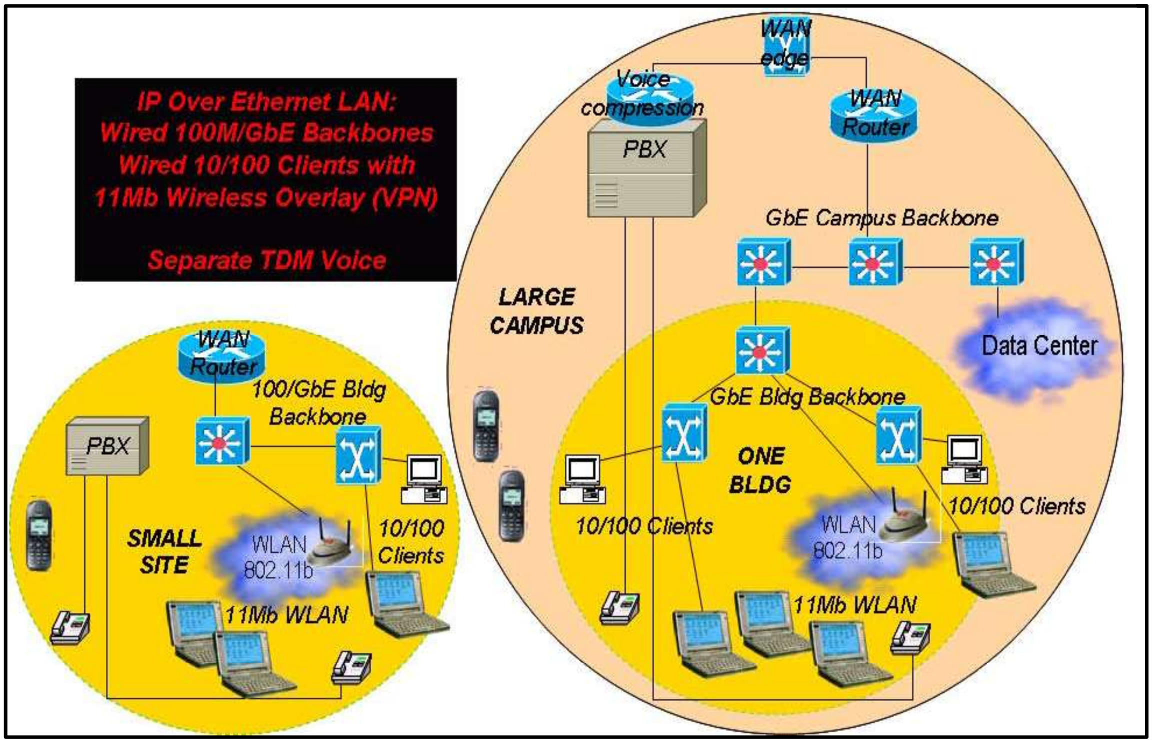

Figure 1: Existing separate LAN, Voice, and WLAN

In order to support voice and video, which are delay sensitive services, data services networks have to address and support basic scheduling capabilities. Since voice is considered by most users and the regulatory domains to be a more critical service than data, securing voice is an important factor. One of the biggest security problems in the enterprise network is the standard "permit all" paradigm of the LAN. While this openness was a catalyst to the growth of computer networks, it also allows devices with security issues to freely connect to the network and potentially compromise other devices. Traditionally LANs have also paid no attention to admission control and cannot separate users with different access rights such as company employees and visitors. As more and more mission-critical services are added to the LAN, steps must be taken to enforce those rights through comprehensive security and QoS mechanisms.

LAN READINESS FOR CONVERGED COMMUNICATION

The LAN infrastructure needs to be robust and redundant in order to support converged services. A typical LAN three-tier architecture is shown in Figure 2. In order to get LAN ready for data, voice, and video convergence over LAN and WLAN, separate changes need to happen at all tiers of the architecture. In the next section we describe required changes in the area of security, power over Ethernet (PoE), QoS, and WLAN integration.

Security

As stated, so far voice and data networks have addressed security by keeping two separate networks with minimum overlap in traffic. However, in converged network environments we are exposing critical services such as VoIP to data network vulnerability. Therefore, during the development of converged networks, all measures should be taken to protect and minimize the impact on data networks. Figure 2 shows the existing three-tier LAN architecture [1, 2]. The edge of the network (access layer) does not have any security deployed allowing anybody to connect to the Intranet. Most of the LAN infrastructure is not ready for converged communication. During a crisis, access-control lists are deployed at the building distribution level to protect from Malware.

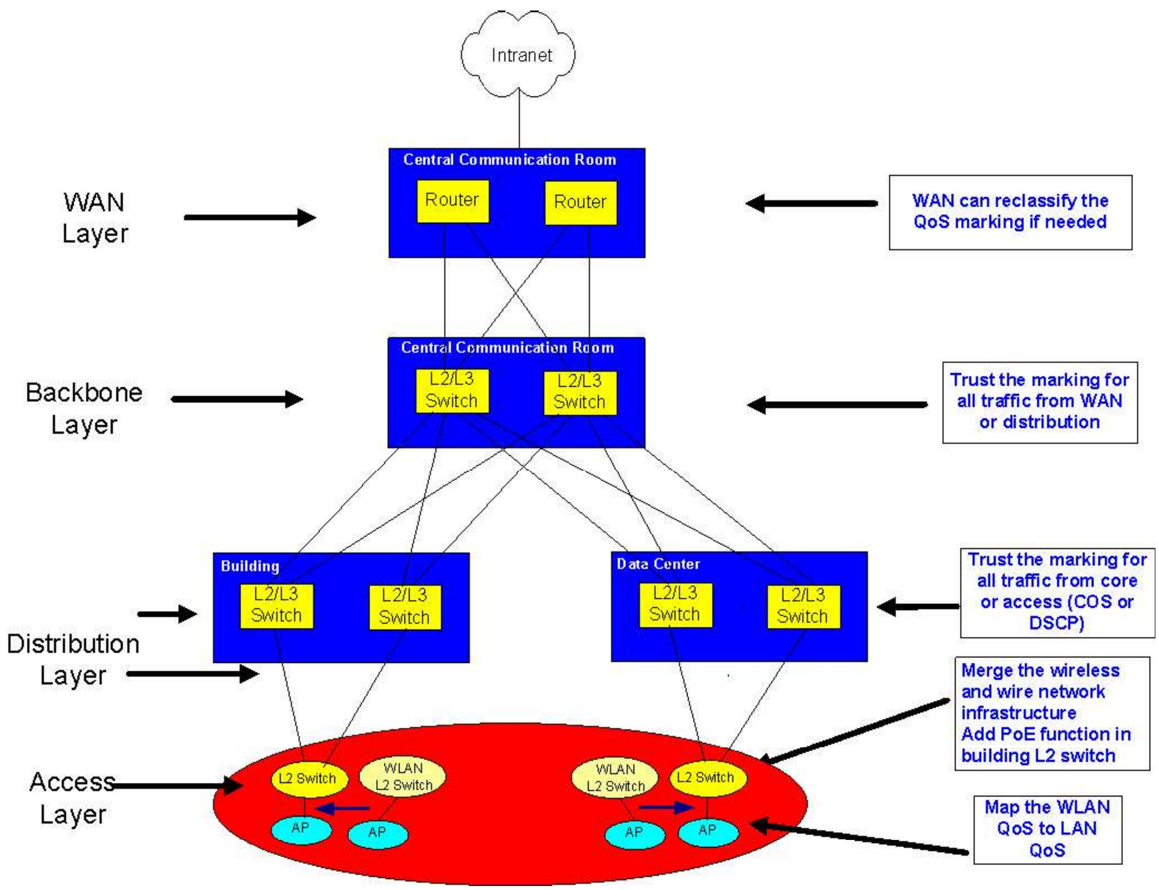

Figure 2: Existing three-tier LAN architecture

Based upon the quantity of building routers, the propagation of ACL can take several hours and during this time sensitive services such as voice may be impacted. In order to avoid this from happening a number of steps should be taken: only authorized machines should be allowed to be connected to the LAN and the ability to detect and remove offending devices at the access location must be developed. In this area, multiple capabilities and technologies need to be leveraged to protect the LAN and reduce the risk of failures due to malicious code. The IEEE 802.1X standard is one example that offers both wired and wireless devices a method to authenticate both the device and the user before allowing access to the network. Based on the Extended Authentication Protocol (EAP), the 802.1X standard allows direct communications using EAP between the end device and a backend authentication (RADIUS) server prior to allowing network access [3, 4]. Only authenticated users and devices are allowed to connect to the production VLAN. All other devices are either not allowed connection to the enterprise or can be placed on limited access networks. Since EAP is by design extensible, it can be expanded to include checking compliance with corporate policies such as operating system patches and virus signatures. When the device does not have the right credentials, it can be redirected to a different network where it can get patched or be limited to accessing the Internet only. This will guarantee that only trusted machines with the right credentials and posture (domain account, operating system level, patch level, configuration level, etc.) can access the

Intranet, a feature that will enhance the security level for both voice- and data-using devices. In the future, stateful level inspection, network intrusion protection systems, or NBAR can also be performed at the distribution layer. This can provide protection from application-level attacks. The distribution layer should continue to be used as the first multi-layered defense with the access layer used to enforce connection policies and connection termination capabilities. Sub/Super VLAN (also known as private VLAN) technology can also be used to isolate some systems within a broadcast domain.

Special attention should also be given to keep all the critical services (DNS/DHCP/tftp, etc.) in separate protected domains or enclaves.

Power over Ethernet

With the advent of the 802.3af standard (Power over Ethernet, PoE) the number of cables used is decreased since power is supplied to the end device over the LAN connection. Moreover, if the edge device supports 802.1q trunking, this single connection can support both a dedicated VoIP device (hardphone) and another (data) device such as a desktop computer or a laptop computer connected through it. A single Ethernet port will thereby support both voice and data devices, and the 802.1q protocol will separate voice from data. It will also require that all access switches support PoE function. Enterprises should make the decision on PoE vs. non-PoE at access switches, based upon a return on investment type of analysis. In an existing building with the right switch it is not always necessary to have PoE-based line cards. Most VoIP hardphones can also use an external power supply that can be connected to a regular building power source. Therefore it is good practice to build all new networks with PoE but for existing buildings, using an existing external power supply is still a very cost-effective option.

Quality of Service

The main objective of QoS within LANs and WLANs is the prioritization of traffic during congestion. Since all the LAN traffic is bursty in nature, it can cause buffer (especially transmit buffer) over-runs and under-runs. The first step in a QoS is to identify the traffic and classify it to enable different traffic types to be processed differently. Typically, access control lists are used to identify the traffic using the source/destination IP address and TCP or UDP ports at Layer 4 or the application signature at Layer 7. Policing or shaping of the traffic can happen at the same device, or alternatively the packets may be marked with a specific priority at Class-of-Service (CoS) bits at Layer 2, Type-of-Service (ToS), or Differentiated-Service Code Point (DSCP) at Layer 3, and those markings can be used later on by other devices. To keep the QoS end-to-end, all intermediate routing/switching devices must trust the marked traffic to minimize the re-marking process. It is also best to identify and mark the traffic closest to the source (normally at the access layer switch in the wiring closet as shown in Figure 2) [5, 6]. Since most of the marking within the Intel LAN environment will be done by the applications running on a system or hardphone, it is essential to allow trust of the endpoint marking within the LAN and WLAN. It is also important to have separate queues for voice (latency sensitive) traffic and for other traffic, and a priority-scheduling scheme should be given to it. In the coming years, new applications/services are being planned to be deployed to improve the productivity of users. VoIP in LAN and WLAN is one of them. Convergence of multiple communication methods will also drive the need for QoS in LANs and WLANs.

WLAN Integration

Many enterprises treat WLAN as unsecure and hence require users to use VPN services before gaining access to corporate resources (Figure 3). This has been the general policy due to the well documented weak security of the Wireless Equivalent Privacy (WEP) security measure. With the latest development in WLAN security standards (802.11i, using WPA2 with authentication based on 802.1X authentication and AES encryption), however, the VPN component can be removed from the wireless network allowing the Access Point (AP) to directly pass traffic to the LAN infrastructure. Recent LAN security-related protocols, mainly MACsec (802.1ae) and MAC key security (802.1af), make the authentication and encryption schemes of WLAN and LAN converge.

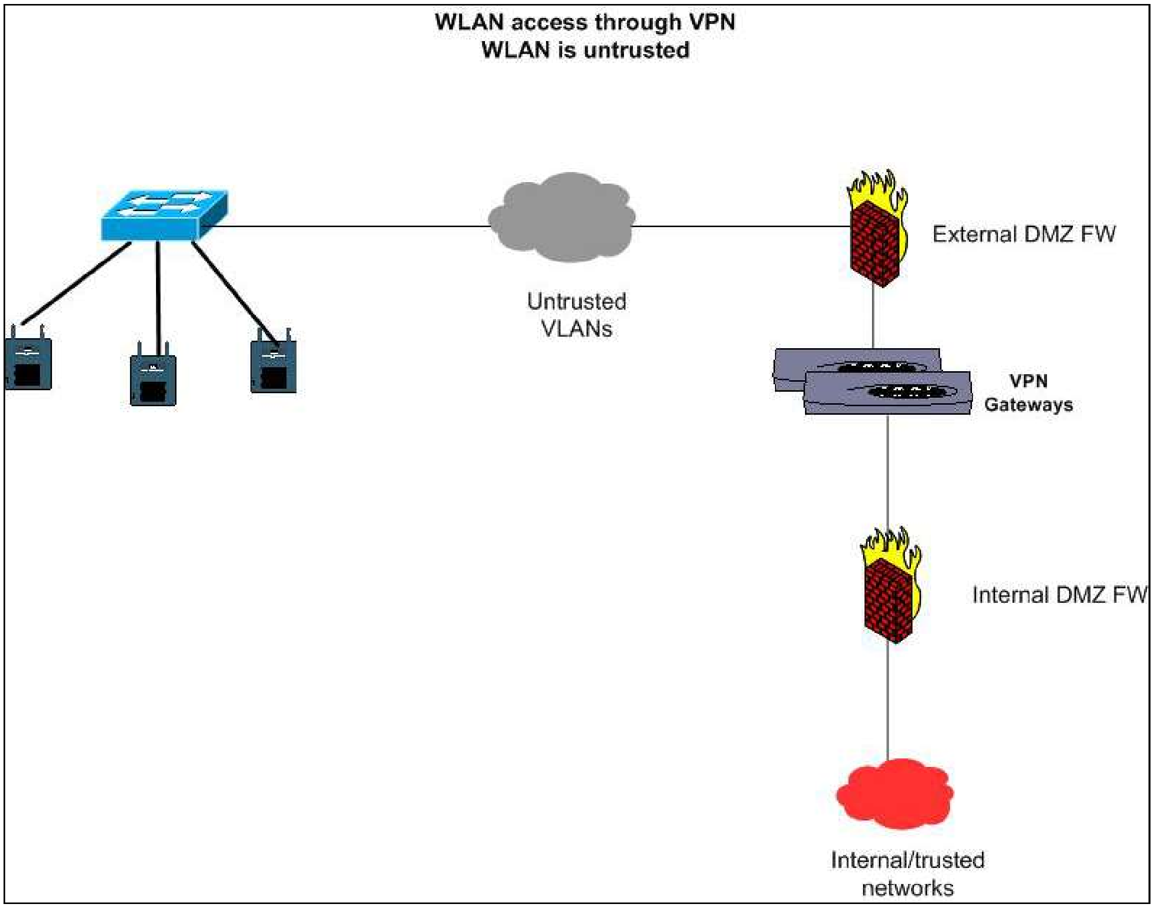

Figure 3: Existing WLAN architecture

New Converged Architecture

Figure 4 shows the integrated converged network where the WLAN is considered as an extension of the LAN, and the endpoint registers with the WLAN controller to provide services. As WLAN access technology changes from 802.11a/b/g to 802.11n, the same architecture can be used to support the new access technology providing enhanced throughput. VoIP becomes the primary voice technology within the building, and it connects to the legacy PBX to provide backward connectivity. However, all VoIP end nodes talk to each other, directly connected by the VoIP server.

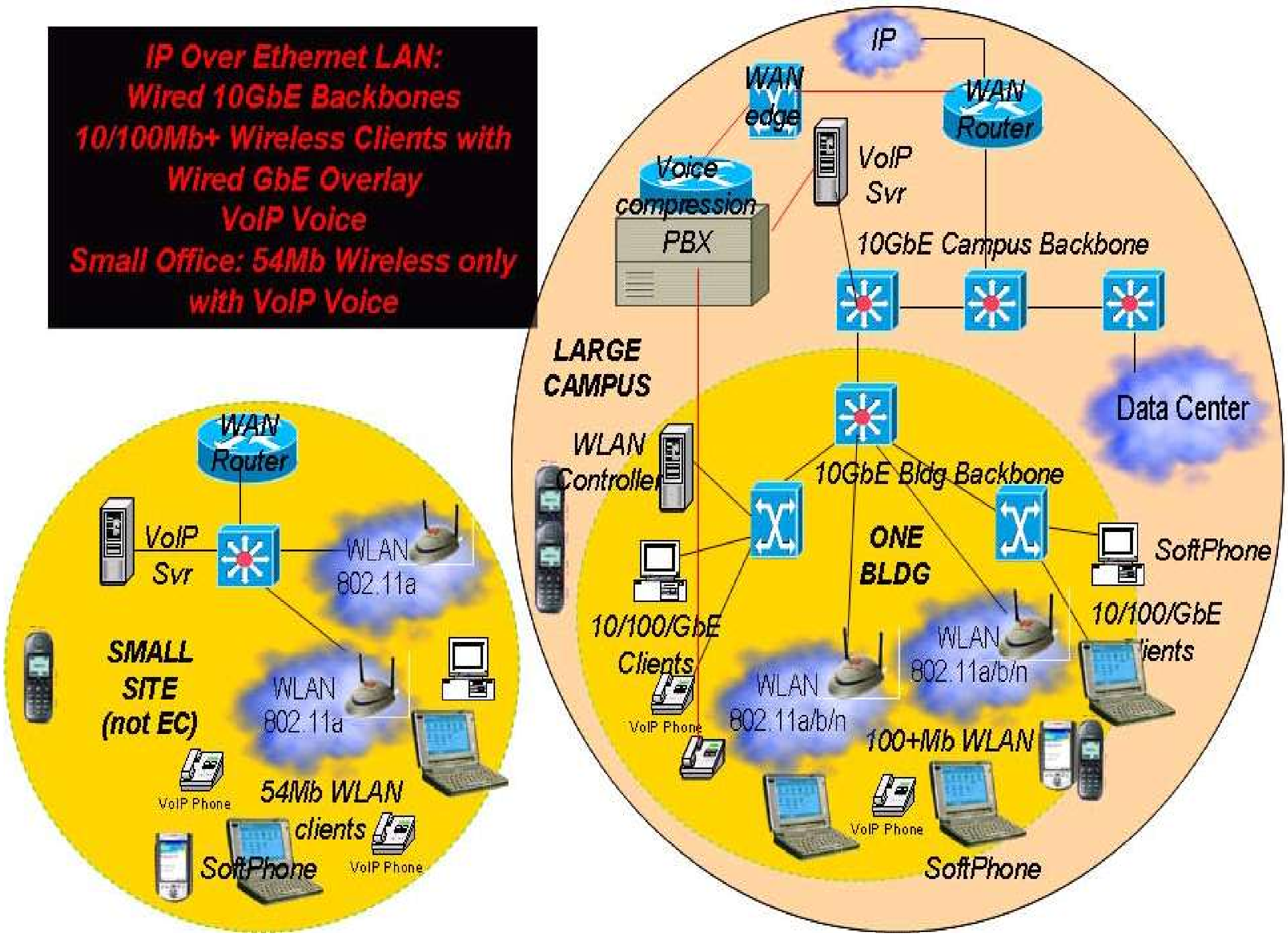

Figure 4: Converged network architecture

FOLSOM LAN AND VOICE CONVERGED NETWORK

Intel is working to build a first converged network at one of its campuses. In this network all 300 users will have only one network connection to their VoIP hardphone and their personal computer will be connected to the VoIP phone. Hardphones will receive power via PoE, and QoS will be used to put the voice traffic on a strict priority queue. All the WLAN APs will be connected to the same LAN switch to eliminate the separate physical infrastructure. Figure 5 shows the high-level topology of the converged network. Before, this site would have had two parallel networks, one for voice and one for data. With this topology, users will only have one connection for all services.

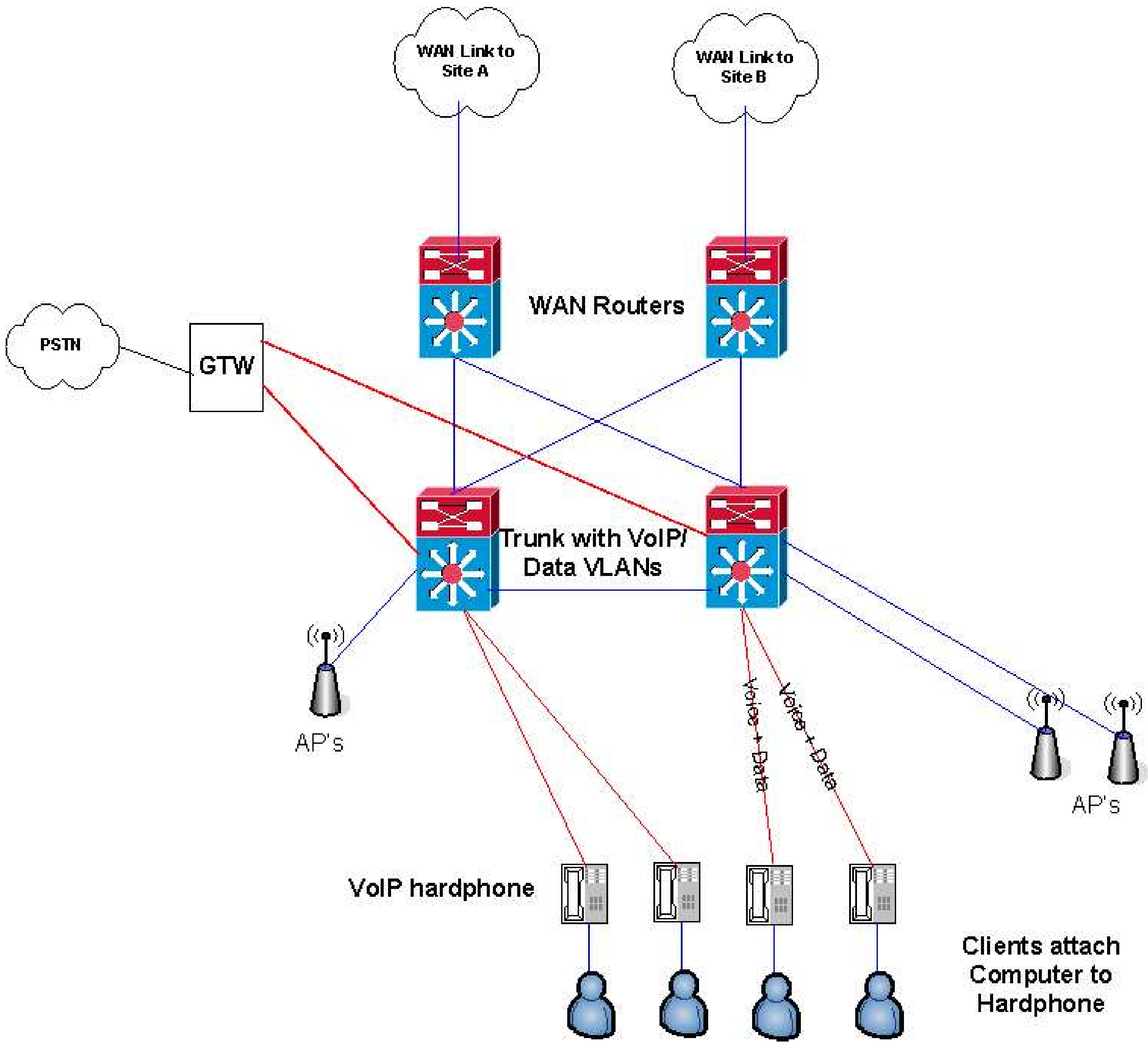

Figure 5: Trial converged network

CHALLENGES

As enterprises strive to converge voice, video, and data over LANs and WLANs, they will face multiple challenges. Some of those challenges are technical, and a significant number of them are business related.

Technical Challenges

The main technical challenges involved in building converged networks are as follows:

1. *New voice backend*–Moving voice from a circuit-based network to a packet-based network involves a radical change in the voice backend. While traditional PBX manufacturers support such a change, a lot of companies can take the opportunity to start fresh with a native VoIP solution. The move should affect user services as little as possible.

2. *Quality-of-Service*–Introducing different services with different network characteristics into the same network requires the ability to prioritize more time-

critical network users such as VoIP over less time-dependent services such as data. Introduction of QoS involves a significant technical challenge and may require a network equipment update and refresh.

3. *Security*–As different networks merge into a single infrastructure, the separation offered by different transports and cabling is lost. In traditional voice networks the only way to access other users' traffic is to physically gain access to their phone cabling or break into distribution frames or PBXs, which are normally locked up. With shared networks, if the network is not protected properly, unauthorized access to other users' data (and therefore VoIP) may ensue without physical attachment with methods such as ARP poisoning.

4. *Training*–Support staff are used to working with the legacy equipment and will have to be trained on the new equipment. Since this is a paradigm shift for most support staff, the required training is extensive. Most organizations have separate support personnel for data and voice; with the integration of voice and video on one data network, support should also merge for better end-to-end service quality.

5. *Troubleshooting*–The resulting converged network is more complex than its component parts. Putting more eggs in the same basket always makes it more difficult to find out the root cause of problems. To make the converged network successful, the right management and troubleshooting tools have to be created for the support staff.

Business Challenges

Besides the above, there are multiple business-oriented challenges that should be addressed before converging networks:

1. *Making the business case (Return on Investment, ROI)*–Probably the single biggest challenge for anyone wanting to make the leap is making the case for this move. In today's business environment changes of this magnitude are unlikely to be approved without a hard dollar calculation showing real savings.

2. *End user expectation reset*–Today, end users are accustomed to very high availability of their voice network and in converged networks that expectation should be level set.

RESULTS

Convergence of services in a common network is inevitable. In order to support high-quality service to the customer, the LAN needs to be redesigned to meet delay, jitter, and loose characteristics. We have developed a design that can meet voice, video, and data needs in LAN and at the same allow the WLAN to be merged with the wired LAN. We equipped a few small offices with this design and are working to implement it in a 300-user office. We have been successful in replacing end-of-life products with new VoIP and are therefore positioned for worldwide VoIP down the road.

DISCUSSION

Converged networks offer great promise for converged communications by integrating voice, video, and data on LAN and mobile networks. As such, converged networks show great promise for the enterprise. However, the challenges are considerable in the areas of business, finance, and technology.

Should enterprises today take the plunge? We believe there is no simple answer to this question and the important thing to remember is that "one size does not fit all" anymore. In deciding whether converged networks are right for an enterprise, we offer the following advice:

- *Core vs. edge*–Core services are normally centralized and controlled more closely than the edge; therefore, starting the move to convergence at the core is sometimes easier. Moreover, a lot of the immediate financial benefits are more easily seen on the core, since long-distance networks are a major part of the network discretionary spending, and convergence at the core can reduce spending considerably. On the other hand, the core is one of the most sensitive and mission-critical environments, if not *the* most.

- *New/Greenfield vs. legacy sites*–When going into new construction an enterprise has to think ahead at least three years. We believe by this time converged networks are going to be the rule rather than the exception. Therefore, we would recommend considering using converged networks at new sites and campuses. Making the case for retrofitting existing infrastructures is much harder.

- *Risk taking*–Making major changes to an existing environment is risky. In making the decision one has to weigh the risk vs. the gain. Starting convergence in a mission-critical manufacturing plant, design center, or customer support center is quite different from making that move in a standard office environment. Start small and expand as you gain experience and confidence.

- *Standard based vs. proprietary*–Using standards offer a much better chance of interoperability, but some of the standards may still be in the certification track. Using existing proven proprietary protocols may be required as an interim step. The timetable for implementation will determine this choice ultimately.

- *Network management*–Being successful in converged networks requires new capabilities for managing the converged networks including QoS management, security management, and real-time troubleshooting tools. Having this kind of management is crucial to success in this complex task.
- *Staff and technical expertise*–Moving to converged networks requires a lot of talent and knowledge, radically different talent and knowledge gained from working with legacy networks. Making the change requires paying close attention to skill levels and training for both design and operations personnel.

CONCLUSION

The communication industry has now widely accepted IP as the universal transport protocol for the enterprise. In the last decade all other transport protocols have converged to IP. Now there is rapid adoption and migration by telephone, cable, and media vendors and industry to move to IP to take advantage of converged networks. This has had a snowball affect within the enterprise: vendors are forcing corning enterprises to go to IP transport for all services. Therefore, the enterprise should start planning to make the appropriate changes to their network to position themselves for new IP services. We have started rolling out converged networks for our small offices and are aggressively working to roll out our first big-size office with a converged service network.

ACKNOWLEDGMENTS

We thank Justin Richardson for piloting the converged network architecture and providing input. We also thank our VoIP LAN working group members: Neil Wallace, Tim Verrall, Shelby Siegel, Dave Lizotte, Charles Weaver, John Gresham, Blaine Bauer, Bob Wasserman, Kevin Heine, and Dave Brooks. We also appreciate the contribution of the ITJ editorial staff.

REFERENCES

[1] "Internetwork Design Guide," Cisco Systems, http://www.cisco.com/univercd/cc/td/doc/cisintwk/idg4/*

[2] "Open System Interconnection Reference Model," Cisco Systems, http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/introint.htm*

[3] "Port Based network Access Control," http://standards.ieee.org/getieee802/download/802.1X-2001.pdf*

[4] "RADIUS Protocol Security and Best Practices," Microsoft Corporation, http://www.microsoft.com/windows2000/techinfo/administration/radius.asp*

[5] "IEEE 802.1p: LAN Layer 2 QoS/CoS Protocol for Traffic Prioritization," http://www.javvin.com/protocol8021P.html*

[6] "IEEE 802.11e-2005 wireless network QoS specification," http://standards.ieee.org/reading/ieee/std/lanman/restricted/802.11e-2005.pdf*

AUTHORS' BIOGRAPHIES

Sanjay Rungta is a principal engineer with Intel's Information Services and Technology group. He received his B.S.E.E. degree from Western New England College and his M.S. degree from Purdue University in 1991 and 1993, respectively. He is the lead architect and designer for the Local Area Network for Intel. He has over 13 years of network engineering experience with three years of experience in Internet Web hosting. He holds one United States patent and two pending in the area of Network Engineering. His e-mail is sanjay.rungta at intel.com.

Omer Ben-Shalom is a lead engineer for Intel's IT Wireless LAN team. He received his B.Sc. degree in Physics and Computer Science from the Hebrew University in Jerusalem in 1996, and his M.Sc. degree in Computer Science from Tel-Aviv University in 2002. He is a lead engineer in the Information Services and Technology Mobility Services group. His e-mail is omer.ben-shalom at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

Session Initiation Protocol (SIP) Evolution in Converged Communications

Mohammad Kolbehdari, Digital Enterprise Group, Intel Corporation
Dave Lizotte, Information Technology, Intel Corporation
Glen Shires, Digital Enterprise Group, Intel Corporation
Scott Trevor, Information Technology, Intel Corporation

Index words: VoIP, Presence, TDM, PSTN, IP PBX, SIP, SIMPLE, RTP, Converged Communications, Interoperability, Security, Enterprise Networking, Mobility, IETF

ABSTRACT

With the telephony business world focusing on Voice over IP (VoIP), numerous IP Private Branch Exchange (PBX) vendors are offering rich new product lines supporting VoIP capabilities, presence integration, and other enhanced multimedia and location-based services. In order to achieve true convergence of these technologies, PBX, phone and network vendors will need to design and support products that will interoperate with each other based upon industry standards. Session Initiation Protocol (SIP) is being widely adopted in the industry [1] as a signaling protocol because it enables data and voice convergence for devices and applications across a wide range of industry sectors. SIP enables voice, video, Instant Messaging (IM) and other media, and facilitates presence and location-based services. SIP's extensibility and versatility enable rapid innovation of new, rich features, and rapid deployment, and it has become a market enabler for VoIP PBX's and IP telephony devices and applications. However, SIP's extensibility has also introduced interoperability challenges as vendors differentiate by extending beyond baseline SIP specifications. Implementation of rich features in a standard, interoperable manner requires ongoing standardization and industry consensus.

INTRODUCTION

Converged communication improves person-to-person business communication with the unification of numerous messaging modalities including voice, video, IM and Presence, and integration with data applications and collaboration tools.

SIP is emerging as a key enabler of converged communication offering an interoperable protocol for such requirements and flexibility for many services' and vendors' networks. SIP differs from other communication protocols by its strong industry support, multi-vendor integration at the application layer, modularity, and common standards. Since SIP is an application-layer protocol, it is transparent of the underlying data link layer topology, which simplifies its deployment.

In this paper we focus on the emerging capabilities of SIP within real-time communication technologies while addressing challenges within the areas of interoperability, security, and enterprise network integration. We describe how the seamless integration of presence-based SIP, VoIP, Mobile IP, SIP mobility support, unified communications, and applications can lead to converged communications.

SIP AND CONVERGED COMMUNICATION

Most legacy telephony devices are dumb endpoints with no memory or processors built into them, as all of the intelligence was located within the proprietary PBX processor software. In contrast, with VoIP PBX solutions, all endpoints have processors and memory thus creating superior intelligence out at the network edge.

SIP is a signaling protocol for the establishment of communication sessions between these smart endpoints. It is commonly used to initiate voice, video, and IM sessions and can also be used to convey presence, location, and other information. SIP has emerged as a key protocol with strong industry support for the deployment of IP-based telephony. In addition to the rich media session and information that it can convey, SIP offers these additional benefits:

- Converged Network: Using a single network for voice and data reduces cost and simplifies management.

- Mobility: SIP provides a user with a logical identity regardless of the device type he is currently using or the device's physical location. This allows users to roam and to switch between devices (such as from a handheld to a computer SIP phone), while remaining reachable through a single address: callers do not need to try numerous phone numbers.
- Enhanced Audio Quality: The traditional public telephone network (PSTN) only transmits a small portion (200 Hz to 3.4 kHz) of the full range of human speech (80 Hz to 10 kHz). Its limited frequency response and dynamic range is the reason people "sound different" over the phone, and must resort to saying "S" as in "Sam," "F" as in "Frank." In contrast, VoIP phones can use a wideband codec to capture and reproduce audio with a much wider frequency response (for example, 50 Hz to 7 kHz) and dynamic range, resulting in a dramatically clearer call with minimal distortion of speech.
- Integrated Presence: A communications session is commonly initiated after identifying the user's availability or willingness to communicate. The publication of the user's presence information can also be used to determine the appropriate type of session to initiate (for example: while in a meeting, the user may prefer to receive an IM, whereas while driving the user may prefer voice).

Interoperability and Standardization

The full potential of VoIP can only be realized if calls are connected over an IP network end-to-end, rather than relying on gateways through the PSTN to VoIP networks. Not only do these gateways add cost and latency, but they block the rich features that VoIP can provide, features that include enhanced audio quality, video, IM, presence, and application sharing. These features require end-to-end IP connections, commonly referred to as direct-IP peering. Interoperability is required for direct-IP peering.

The rapid convergence to SIP is a strong step towards interoperability, but the SIP specification alone is not sufficient. SIP's advantages include simplicity and extensibility. However, its extensibility has resulted in implementations that extend SIP in incompatible and non-interoperable ways. SIP and many related protocols have been developed through the Internet Engineering Task Force (IETF) standards organization.

The primary SIP specification, RFC 3261, specifies how sessions are created, modified, and terminated, and it defines its use with registration and proxy servers. However, many additional IETF specifications are commonly used in conjunction with RFC 3261; for example, additional specifications are used to define how session capabilities are formatted and negotiated, how firewalls and NATs are traversed, and to clearly define certain transition states and call features. Additional specifications are used to define the codecs for various media types and the transport they flow across, typically RFC 3550 Real-time Transport Protocol (RTP). Other specifications are used to define presence and IM features. Numerous other specifications are used to define security, identity, and authorization; with additional specifications defining the underlying transport protocols.

Development of these specifications has been, and will continue to be, done through IETF working groups including the "SIP" working group to develop the primary protocols, the "SIPPING" working group to determine and document new requirements, and the "SIP for IM and Presence Leveraging Extensions" (SIMPLE) working group to define IM and presence applications. As the protocols continue to evolve, they are in various stages of becoming a standard. Some are full Internet Standards; others are in the stable but not yet at the ratified "Request For Comments" (RFC) stage, while others are rapidly evolving Internet Drafts. Specifications developed by other organizations are also commonly used, including some codec specifications developed by International Telecommunication Union (ITU) [5] and some transport protocol specifications developed by the Institute of Electrical and Electronics Engineers (IEEE).

Clearly, consensus throughout the industry is needed for consistent implementation and a rich level of interoperability to be achieved. The greatest roadblock or challenge for an enterprise attempting to provide converged communications seems to occur within the integration of different vendor products into a seamless solution. The lack of interoperability of vendor products can cause a project budget to increase to address the integration efforts. SIP functions as a signaling protocol, but does not support the enhanced capabilities vendors are providing within their products.

Despite lack of ratification, many extensions have been widely implemented in numerous products. These extensions can however compromise the interoperability between vendor's solutions. When it comes to integrating rich presence and other key value-added features between vendor products, it has become clear that SIP interoperability is not enough. Open source implementations of SIP stacks, such as reSIProcate, have helped achieve interoperability among various implementations, but since the stacks are not a complete implementation of a SIP product, they alone are not enough to ensure rich interoperability.

Some industry organizations [7] are beginning to address these concerns. For example, the 3rd Generation Partnership Project (3GPP) has created the IP Multimedia

Subsystem (IMS) specification which is a standardized implementation based on SIP. The GSM Association (GSMA) has a series of interoperability trials based on SIP and IMS. The SIP Forum has begun work on specifying best industry practices for SIP and also coordinates SIP Interoperability test events (SIPit). Any vendor with an implementation of SIP is encouraged to attend these bi-annual events where implementations are tested against each other and problems are often resolved on-site. Although specific vendor results are not made public, SIPit has succeeded in eliminating many of the hurdles of interoperability.

Development and Deployment Benefits of SIP

The SIP specification was originally published in 1999 as IETF RFC-2543 [2] and updated in 2002 as RFC-3261 [3]. Whereas signaling protocols such as H.323 utilize a single administrative domain architecture, SIP can be peer to peer and across domains. By providing an effective communication method between peers, SIP enables innovation of client features without requiring deployment of additional network infrastructure.

SIP utilizes a standard call control mechanism to set up, manage, and tear down a communication session. It is the first protocol that can run over reliable and unreliable transport protocols, has request routing capabilities for performance and control, and is extensible. Carried within the SIP message body, SIP utilizes the Session Description Protocol (SDP) to describe the session parameters such as call attributes, Real-time Transport Protocol (RTP) and payload format, and User Datagram Protocol (UDP) port selection, while also negotiating and exchanging media capabilities such as audio codec selection, video, or shared applications. A typical session initiated by SIP is a packet stream of the RTP, a standardized packet format for delivering audio and video over IP. SIP defines basic transactions and is extensible, scalable, and allows for supplementary information to be carried within the payload allowing devices to make intelligent call-handling decisions and invoke other application-level services such as IM and Presence. SIP is the first protocol to enable multi-user sessions regardless of the media content.

SIP is similar to the Hyper Text Transport Protocol (HTTP) in the way that messages are constructed. This allows developers to easily create SIP applications using common programming languages and Web services, to re-use code and tools, and to more easily debug applications.

SIP Re-uses Existing Internet Features

Prior to SIP, a typical VoIP PBX used signaling protocols such as H.323, H.245, and H.225 for the call setup, control, and teardown of a voice call. With IP telephony, the H.323 protocol suite had to be revised, as the absence of a standard for VoIP resulted in incompatible products. Only a portion of the H.323 architecture is used for VoIP when it comes to audio calls. One of the drawbacks of H.323 is that it will first establish the session then negotiate the capabilities and features for that session. The H.323 protocol provides only a numbering scheme for identities or addresses, thus does not provide the scalability and flexibility of the more versatile URI-based addressing. SIP's URI-based addressing allows callers to use either the URI names (which might be the same as the recipient's e-mail address) or mapping to a numeric dialing plan.

With the emergence of converged communications, the SIP protocol offers that attractiveness of re-using existing Internet features for real-time, mobile, and seamless collaboration. SIP uses a large selection of protocols that are already being utilized by applications for the Web, Internet, and IP-based networks. IP networks route differently than traditional PSTN telephony networks. The basics of IP routing is to route a packet to a desired destination or intermediate point that can make further routing decisions based upon the final destination's IP address. Because a user typically does not know the IP address of the end user they are trying to communicate with, the use of a Domain Name System (DNS) is utilized. Utilizing VoIP telephony with SIP, a user's identity is defined by a Uniform Resource Identifier (URI) based upon his/her IP address, username or phone number, and host name, and not on a distinct telephone number tied to only one location, as is done in traditional telephone systems. SIP uses DNS procedures to resolve a SIP URI and locate the appropriate SIP registrar for the call recipient. Based upon these services end users can utilize one identity name and become reachable anywhere on the network upon which they reside.

When a user wants to place a call, a SIP invite is initiated. The SIP communication session will determine the end device to be contacted and the user's availability and willingness to communicate. The session will also negotiate the media and other capabilities, such as the audio codec to be utilized and may even renegotiate any additional features or capabilities needed during the session. SIP also provides the ability for either endpoint (or an intermediate proxy) to tear down and terminate the call.

SIP is utilized in various architectural components including User Agents (UA endpoints), registrars, proxy, and redirect servers. A SIP registrar allows for all SIP user agents to register and authenticate to the network as an active user capable of placing calls. The registrar also acts as a repository for SIP URL/URI's and other identity information. A SIP proxy server can perform application-level routing of SIP requests and SIP responses for the requested home location services. If a proxy cannot identify the request it will send it to a redirect server. The redirect server does not forward SIP requests but points the proxy to contact another server that might know where the requested INVITE or UA resides. When utilizing a non-SIP endpoint, such as a legacy digital or analog device, the call will utilize a SIP gateway to act like a UA to allow for the protocol translation to non-SIP networks such as H.323, Media Gateway Control Protocol (MGCP), and Public Switched Telephone Networks (PSTNs).

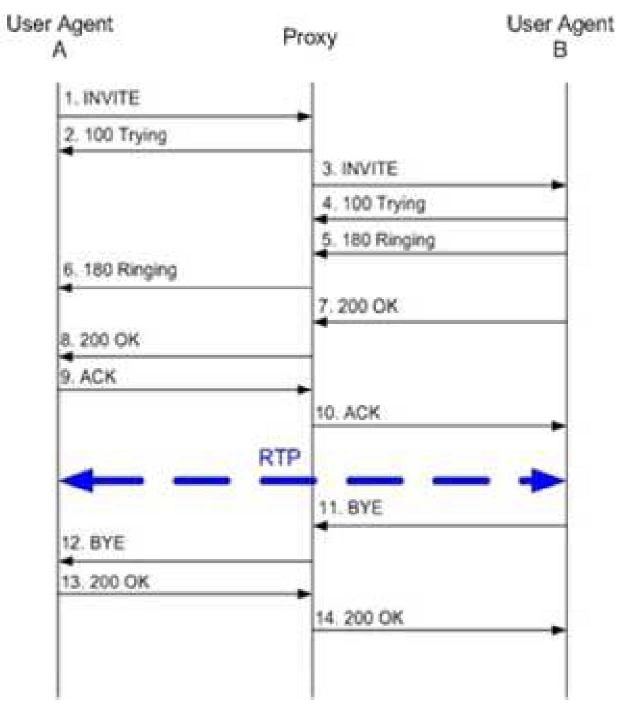

Figure 1: SIP call between user agents

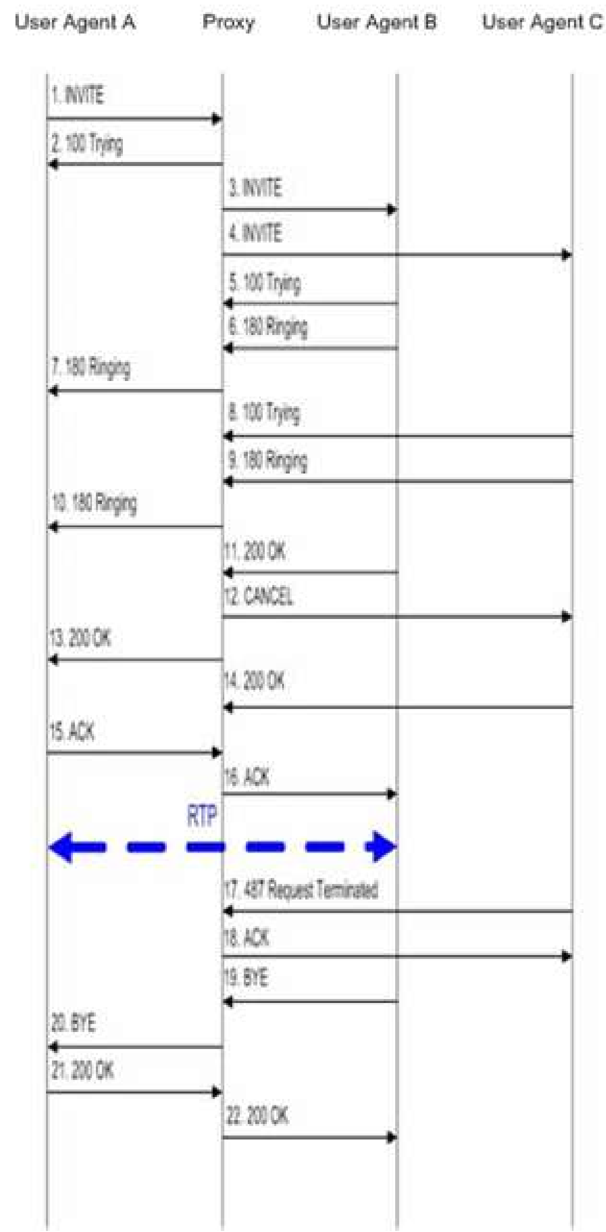

Figure 2: Multiple user agents and forking

Mobility

Mobility allows devices to stay connected even as the device moves between networks. Two methods that implement VoIP mobility are Mobile IP and SIP Mobility. Mobile IP is a network layer approach: it operates below the application layer and is applicable to most applications including VoIP applications based on H.323 and SIP, but it can add latency due to tunneling of the data stream. SIP mobility uses application layers (3 or 4) and augments existing VoIPs such as SIP or H.323 [5]. Being an application layer protocol enables SIP mobility to be deployed easily without requiring the network infrastructure support that Mobile IP requires. In addition, SIP mobility enables not just device mobility, but also personal mobility, allowing a user to easily switch between different SIP devices.

Mobile IP allows applications to use a given IP address and stay connected to devices regardless of their locations. When users with mobile devices leave the network that their device is associated with and call their home network and enter the domain of a foreign network, the foreign network uses the Mobile IP (IETF, RFC 3344) [4] to inform the home network of a Care-of-Address (CoA) to which all packets for the user's device should be sent. The Mobile IP network topology is shown in Figure 3.

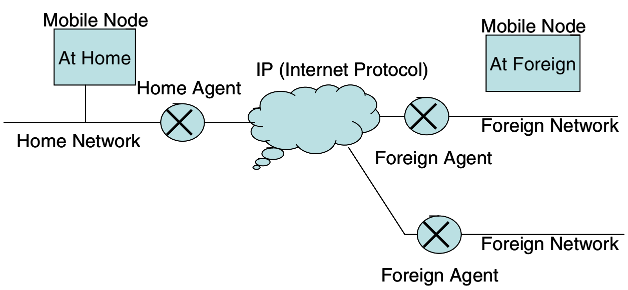

Figure 3: Typical Mobile IP topology

IP mobility is supported by transparently binding the home address of the mobile node with its CoA which is the termination point of the tunnel toward the mobile node when it is not in the home network. IP mobility binding is maintained by specialized routers known as mobile agents. Mobile IP has the following three main components: mobile node, home agent (HA), and foreign agent (FA). The three main phases of the Mobile IP process are as follows:

- Agent discovery–a mobile node discovers its FA and HA during agent discovery.
- Registration–the mobile node registers its current location with the FA and HA during registration.
- Tunneling process–a reciprocal tunnel is set up by the HA to the CoA to route packets to the mobile node as it roams.

The CoA is a temporary address that is valid while the mobile node is attached to the foreign network domain. In Mobile IP Wireless Local Area Networks (WLANs), the two mobility agents, the HA and the FA, coordinate, update, and authorize the connections and CoAs for clients from foreign networks. These connections are provided by binding the update message sent by the HA to the corresponding node. The bound message allows VoIP traffic and messages to be directly tunneled between the caller node and the mobile node. There are two key issues that can arise: the first is when roaming occurs between two foreign networks while a call is in progress between a caller node and a mobile node, and the second one is the timing requirement that must be predefined for WLANs or handset designers. The timing requirements are the period of time needed for a station to associate with an Access Point (AP), the period of time needed by a handset to associate with a foreign network, the period of time to bind to a foreign network and create a new CoA, the period of time needed to send packets directly between the Mobile Node and a Caller Node, and the period of time needed to bind update messages from an old foreign network to a new foreign network.

SIP mobility (IETF, RFC 3261) [3] supports mobility for VoIP applications by providing handoff capabilities at the application layer. The SIP mobility support protocol uses the concept of a Visited Registrar (VR) in the foreign networks. The SIP mobility support with VR features combines some of the functions of a SIP proxy server, location server, and user agent. The SIP proxy server enables SIP [6] to handle both firewall functions and Network Address Translation (NAT). SIP is designed to support roaming so that a user can be found independent of the device he/she is using and its network location. For example, with SIP, a call on a handheld phone can be transferred to a computer SIP phone. The SIP mobility approach at the micro-mobility implementation level is very similar to the concepts of foreign network and home network. In SIP mobility support, the FA of Mobile IPs is replaced by a SIP VR and foreign network. The Mobile IP HA is replaced by a combination of a SIP proxy server, a location server, and a user agent server. Advantages of SIP mobility support are of using the existing IP-based network without modification as well as being fully supported by the Windows* environment (Windows XP*) making possible a rapid deployment in the market place. Figure 4 shows VoIP and IP-based network configuration.

* Other names and brands may be claimed as the property of others.

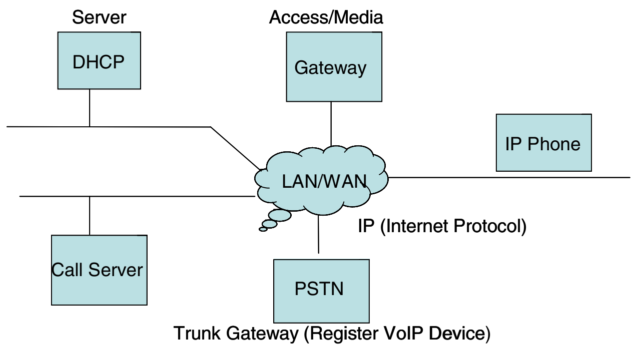

Figure 4: VoIP and IP-based LAN/WAN topology

VoIP Security

Because VoIP uses smart endpoints over a shared network, it provides many benefits and features that are not available with legacy phones over traditional PSTNs. With this flexible VoIP architecture comes additional concerns about security and privacy. However, proper system design can address these issues and in fact, make VoIP more protected, private, and manageable than the PSTN.

Smart endpoints provide a powerful tool to address security and privacy concerns. The traffic can be encrypted at the endpoints and throughout the network. Media traffic is commonly encrypted using Secure RTP (SRTP), and SIP signaling traffic is commonly encrypted using TLS (SSL) and S/MIME. All network traffic can be further encrypted using IPSEC Encapsulating Security Payload (ESP).

To store the protected encryption keys for these protocols, smart endpoints can also provide protected storage mechanisms such as the Trusted Platform Module (TPM) included in many PC platforms. Secure protocols are used for key distribution, such as the Multimedia Internet KEYing (MIKEY) and the Internet Security Association and Key Management Protocol (ISAKMP).

Smart endpoints can also provide identity authentication and attestation. Whereas traditionally identity was provided only through physical means (anyone who could physically access the phone or network could impersonate another caller), smart endpoints can provide an additional level of user authentication through the use of passwords, biometrics, or other means. Attestation to this identity can then be communicated to remote users.

CHALLENGES AND RESPONSES

VoIP in the Enterprise

VoIP technology has matured, and interoperability issues are being resolved. High-bandwidth wireless LANs and WANs have extended access to VoIP enabling more personalized and better integrated services at a reduced cost. However, for VoIP to fully deliver rich, reliable, cost-effective services, the major challenge will be to enhance the interoperability, security, QoS, and bandwidth management for wide-scale VoIP deployment. The desired goal of enterprise deployments is to deploy VoIP to that greater than 90% of enterprises by the next few years. The growing demands are an easily deployable, high-quality VoIP solution, interoperable across network and PBX environments, delivering a new level of performance to the VoIP marketplace, and integration with IP phone-based PBX technology. In some cases, PBX equipment is modular and supports VoIP interfaces such as Line Replacement Units (LRUs). The VoIP gateway is introduced to integrate VoIP into the existing enterprise TDM PBX environment. Enterprises are expected to gradually transition to VoIP because of the expense of legacy analog PBXs and phones and end-of-life of service contracts that support that equipment. There are evolutionary steps that the enterprise can take to move its global telephony environment to a full IP PBX network. A typical TDM PBX voice mail system that is currently adjunct to each distributed PBX can be migrated to the PSTN network or transitioned to an IP-based network by using SIP-based solutions. A VoIP PBX gateway and media server can be provided within the legacy PBX distributed environment to support both TDM and IP-based endpoints, thus providing a large investment protection during the transition period to a standardized IP-based telephony approach and design. New infrastructure sites can be planned to support full IP PBX solution designs with standardized call control capability. In addition, a standardized call control architecture can be developed to support the minimization of numerous legacy PBXs to a handful of centralized PBX media servers. This design method will be key, supporting zones or regions based upon the infrastructure architecture of the data network topology.

There are three critical performance issues that need to be focused on for VoIP deployment: latency–the end-to-end delay; jitter–the variable delays in each voice packet; and packet loss–the dropping of individual packets caused by network congestion. Specific values must be reached in order to ensure that the user has an experience that is the same or better than using TDM telephony. The latency, jitter, and packet loss of the VoIP system are issues regardless of the application of the network data technology. VoIP traffic is very sensitive to dropped packets, network latency, jitter, and packet loss. The key to success for reliable VoIP systems is to control those three major issues. Acceptable VoIP quality requires a latency or delay of not more than approximately 300 ms. Jitter causes irregularities in the flow and delivery of data, and although most vendors have successfully solved this issue using jitter buffers to smooth out the delivery of voice packets, excessive jitter can cause significant additional latency. While slight packet loss is typically not noticeable by users, significant packet loss results in moments of dropped audio and excessive packet loss can cause dropped calls. Queuing priorities solve many jitter and packet loss problems by ensuring timely delivery of voice packets using prioritized-packet method when the reprioritize packets are ordered on the IP network.

Network Capacity and WLAN

Seamless communication is increasingly demanded between data, voice, and video media. Whereas legacy communication technologies use a separate telephony network for voice and an IP network for data, a converged voice and data network provides richer features, integration, and multiple access options at a lower cost. However, a converged network also raises several deployment, configuration, and planning issues concerning QoS, call control, network capacity, provisioning, and architecture. Figure 4 illustrates the concept of IP-based integrated network communications. The network capacity is directly related to the throughput of the network infrastructure as well as the bandwidth requirements for typical voice, data, video, and media applications using IP-based packets. Codecs that offer voice compression help to derive as much capacity as possible by minimizing the packet size. However, to minimize latency, a new VoIP packet is typically sent every 20, 30, or 60 milliseconds, resulting in many small packets that the network must efficiently deliver. WLAN deployment has additional challenges including the variability in number and types of devices connected to an access point, and possibility of radio interference.

Infrastructure Integration

A phased approach can be used to deploy a converged communications environment in a large company. Evolutionary steps will protect current telephony investments while companies migrate to IP-based network solutions. In the interim there will most likely be a mix of older systems using proprietary protocols with newer ones based on SIP. Gateways can provide a smooth migration path for users, but may come at the cost of additional complexity and additional licensing costs. Such gateway services provide the first steps in integrating IM and presence into the telephony environment.

Client Application Integration

True convergence on the client is more than providing all the features in a single application interface; it also requires that they be ubiquitous and available across all the applications on the device; for example, the ability to quickly communicate with the authors of a document one is viewing. Users also demand the ability to quickly and seamlessly move between modes of communication. For example, a user may start out using an IM with a colleague, but then determine that a voice call would serve her better. Later, she might add a data collaboration session to the mix in order to review a document being discussed. Standards-based protocols such as SIP provide the foundation to achieve this level of integration on the client.

Influencing the Vendors

As mentioned earlier in this paper, the desire for vendors to differentiate their products on the basis of features, combined with the relatively immature state of standards makes it difficult to achieve the vision of seamless convergence of real-time collaboration capabilities. Progress is being made with the adoption of SIP by many vendors. More is needed however in several areas including interoperability between wideband audio codecs, rich presence, and direct-peering between SIP networks. Customers have perhaps some of the greatest power to influence vendors. It’s important to push vendors to interoperate. Don’t accept the status quo of “convergence” that is only converged into a single client interface. Look for and encourage vendor solutions that allow for interoperability both on the infrastructure and at the client/device level.

CONCLUSION

With the evolution of SIP as the standard signaling protocol for VoIP telephony, numerous application-level features and capabilities are being developed to advance mobility and productivity for businesses and their end users. Interoperability between various vendor solutions is key to enabling end users to richly communicate through the device type of their choice, regardless of their global location.

SIP is the standard for the establishment of multimedia sessions, including voice, video, and IM; and for conveying presence, location, and other information. SIP-based communications deliver a suite of solutions that can significantly enhance users’ communication options and productivity.

ACKNOWLEDGMENTS

We are grateful for the encouragement and help that we received from our managers, particularly Mark Cooke, Alan Blair, Alex Ahmadian, Mike Powell, Michael Stanford, Jeff Morriss, and Bala Cadambi. Additional thanks go out to individuals who reviewed the paper and provided valuable feedback: Sanjay Rungta, Greg Trusley, Don Meyers, and Duncan Glendinning.

REFERENCES

[1] Regis J. Bates, *Voice-over IP (VoIP)*, McGraw-Hill, Jan. 2002.

[2] Handley, M., Schulzrinne, H., Schooler, E., and Rosenberg, J., "SIP: Session Initiation Protocol," *IETF RFC-2543*, March 1999.

[3] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and Schooler, E., "SIP: Session Initiation Protocol," *IETF RFC-3261*, June 2002.

[4] Perkins, C., "IP Mobility Support for IPv4," *IETF RFC-3344*, 2002.

[5] ITU_T "SG 16 H.323 version 1–Version 4 Standard."

[6] Avaya Technologies, White Paper "Enterprising with SIP."

[7] www.cs.columbia.edu/sip/*

AUTHORS' BIOGRAPHIES

Mohammad Kolbehdari is a senior staff architecture engineer in the Digital Enterprise Group. He received Ph.D., M.Sc., and B.Sc. degrees in Electrical Engineering from Temple University, Philadelphia, Pennsylvania, University of Mississippi, Oxford MS, and Communication University Tehran, Iran, in 1994, 1991, and 1986, respectively. He conducted research in the EM lab at Temple University prior to joining the Department of Electronics, Carleton University, in Ottawa, Canada as a research associate faculty member from 1994 to 1998. He joined Intel in 1998 and has been involved in planning and architecture development teams including those of the Pentium® 4 processor FSB and PCI Express* designs. His areas of interest are EM modeling, Power Delivery Networks, Signal Integrity, Circuit and I/O Serial Differential Design, System-level Architecture, Platform Initiatives, Clock Jitter Characterization and Measurements, and Communications and Networking Architecture. His e-mail is mohammad.kolbehdari at intel.com.

Dave Lizotte is a senior voice engineer at Intel Corporation in Folsom, California. He is currently focused on new capabilities development for Voice over IP and convergence solutions. He has in-depth experience in design and implementation of solutions incorporating voice technologies and related specifications. Dave holds a Telecommunications Engineering degree from the J.M. Perry Technical Institute, WA. He joined Intel in 1994 and can be reached at dave.a.lizotte at intel.com.

Glen Shires is a principal engineer in Digital Enterprise Group Architecture and Planning at Intel. Glen works with cross-company teams to define VoIP and SoIP (Services over Internet Protocol) end-to-end system architecture and usage models. He brings a system-wide approach to architecting platforms, having experience developing products at many levels from logic circuits and board design to architecting and developing software for major embedded and PC products. He specializes in speech, video, telephony, and Internet technologies and has been awarded five patents. His career began at Intel as an applications engineer for the Intel386™ microprocessor. Glen received his B.S. and M.S. degrees in Electrical and Computer Engineering from the University of Wisconsin-Madison. His e-mail is glen.shires at intel.com.

Scott Trevor is a senior engineer/architect within Intel's Information Technology division. Scott's current focus is developing the architecture for Intel's real-time communication capabilities as part of Intel's converged communications program. He specializes in the areas of presence infrastructure and its enablement and integration with various communication media. Scott has been working to architect, engineer, and deploy various collaboration products in the IT industry for 14 years, and has been working for Intel since 1994. Scott received his B.S. degree at Old Dominion University, Norfolk VA. His e-mail is scott.trevor at intel.com.

® ™ Pentium and Intel386 are trademarks or registered trademarks of Intel Corporation of its subsidiaries in the United States and other countries.

* Other names and brands may be claimed as the property of others.

Standards-Based Interoperability for the Advanced Telecom Computing Architecture (AdvancedTCA*)

Steve DeNies, Digital Enterprise Group, Intel Corporation
Jay Gilbert, Digital Enterprise Group, Intel Corporation
Ashok Mishra, Digital Enterprise Group, Intel Corporation
Karel Rasovsky, Digital Enterprise Group, Intel Corporation

Index words: Interoperability, AdvancedTCA, Advanced Mezzanine Card, AMC, Infrastructure, Communications, PICMG, Validation

ABSTRACT

The adoption of Modular Communications Platforms (MCP) and Advanced Telecom Computing Architecture* (AdvancedTCA*) has significantly matured during the past two years and is now achieving broad market acceptance by major Telecom Equipment Manufacturers (TEMs) and Service Providers (SPs). While open standards such as PICMG3.x are firmly in place in the industry, a barrier to broad adoption is the lack of uniform implementations of these standards by vendors. To date, this has been kept manageable through periodic plug-fest events hosted by the PICMG organization, but to achieve truly uniform interoperability the AdvancedTCA ecosystem must coalesce around a set of procedural standards on how to interoperate. Intel is working with fellow travelers on this industry-wide effort to drive ease of integration by TEMs and System Integrators (SIs) while lowering their validation costs and overhead as well as reducing the development cycles experienced by building block suppliers. This ultimately increases the competitiveness of the AdvancedTCA architecture with SPs. It will be achieved through the creation of interoperability requirements driven and owned by the AdvancedTCA industry. In this paper we focus on driving a common understanding in the AdvancedTCA ecosystem around interoperability and the path to achieve it.

INTRODUCTION

Modular Communications Platforms (MCPs) based on standard AdvancedTCA are experiencing rapid adoption across the User Plane, Control Plane, and Application Plane of next-generation communications networks. Leading TEMs are specifically focusing on AdvancedTCA deployment in green field opportunities, such as the build-out of 3G networks and emerging IP Multimedia Subsystem (IMS) network elements. The industry's expectation is that these platforms will be rapidly integrated from best-of-breed modular ecosystem components that offer basic interoperability "out of the box." To help realize this promise, standards-based Commercial Off-The-Shelf (COTS) modular components must be specified, designed, and tested for the all-important attributes of interoperability.

Interoperability Defined

Generally, interoperability is the ability of the components of a system to communicate with each other using standard interfaces and protocols while operating in a common environment. The scope of the interoperability is at the system level, not just at the component level; it does not include the specific functionality of an AdvancedTCA component.

For the purposes of this paper, AdvancedTCA interoperability is meant to emphasize multi-vendor product interoperability. Subjects that represent interoperability points for mechanical, electrical, thermal and management follow:

- Mechanical–Connector alignment, dimensions, front panel characteristics.
- Electrical–Electronic signaling across backplane, clock presentation, fabric options .
- Thermal–Shelf airflow and thermal tolerances including volumetric airflow, board cooling.

---

* Other names and brands may be claimed as the property of others.

- Management–Shelf management, power and cooling management, hot swap and E-keying.

Under the options allowed by PICMG, two boards might well conform to the specification, but not talk to each other. The first step to attaining interoperability is to specify the basic level of conformance to PICMG specifications. This involves defining a subset of the PICMG allowed options and submitting products to formal validation to this subset. The second step to attaining real interoperability amongst different vendors is to validate a product's interoperability with other vendors' products.

Today, there is growing evidence throughout the industry that integrating multi-vendor products to build AdvancedTCA platforms is oftentimes challenging, due to limitations in interoperability of the products involved. We will illustrate these challenges and conflicts using prominent examples and discuss the root causes. The industry lacks comprehensive set of requirements which can be incorporated into product designs. The current approach to interoperability testing is unstructured and it lacks test metrics that is necessary for benchmarking interoperability compliance. If left unchecked, the AdvancedTCA product ecosystem faces likely fragmentation under too many non-interoperable custom Stock Keeping Units (SKUs), possibly missing out on the true standard high-volume potential.

A disciplined, industry-wide approach is required to evolve toward better interoperability of AdvancedTCA products. Intel is driving such an interoperability initiative, aiming to ultimately ease the integration challenges for TEMs and System Integrators (SIs). In this paper we focus on the main pillars behind this initiative and the progress accomplished to date. The first step is collaboration on the interoperability requirements which would lay out a set of testable requirements essential to interoperability. Equally important is the creation of complementary test procedures and underlying tools, that will validate conformance to the interoperability requirements. This lays the foundation for industry-wide interoperability certification of AdvancedTCA products backed by independent test labs.

CURRENT STATE OF ADVANCEDTCA INTEROPERABILITY

The rapidly growing number of AdvancedTCA products from a multitude of suppliers offers an unprecedented choice to platform integrators. The challenge is to ensure that all standards-based AdvancedTCA products will interoperate with each other "out of the box."

Since the inception of PICMG3.0, Intel had spearheaded early efforts to promote interoperability of AdvancedTCA products. Out of industry collaboration came the first Design Guide (DG), a collection of Best Known Methods for designing AdvancedTCA building blocks. As such, the DG is helpful in guiding toward common designs, but falls short of the needed interoperability focus, and is not backed by either test specifications or a certification framework.

The journey toward creating a cohesive set of interoperability requirements starts with understanding the key interoperability challenges. Based on field experiences, the top interoperability issues can be found in the following areas:

- Thermal interoperability of shelf and boards.
- Manageability as demonstrated by shelf/system and shelf/boards interactions.
- Mechanical issues related to boards and shelf.
- Interconnect implementations of various PICMG3.x options.
- Advanced Mezzanine Card (AMC) interaction with AdvancedTCA carrier boards.
- Presenting a common interface paradigm for service personnel.

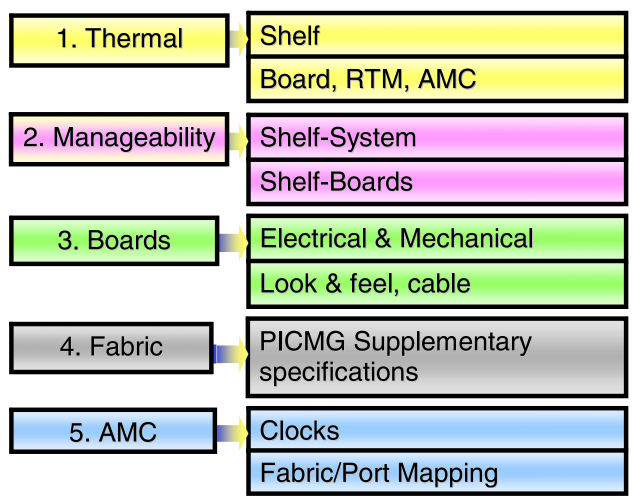

Figure 1: Top interoperability challenges

Challenges in Manageability Area

The AdvancedTCA shelf management system monitors, controls, and ensures proper operation of boards and other shelf components. The key element of this subsystem is the shelf manager that interacts with boards and other shelf resident Field Replaceable Units (FRUs) inside the shelf and also acts as an agent of the system manager, which in most common configurations is located outside of the shelf. From the shelf manager's perspective there are two communication interfaces–one directed internally towards the shelf resident FRUs (boards, etc.) and the other directed externally towards the system manager, as shown in Figure 2.

The PICMG 3.0 base specification defines in great detail, the interactions between the shelf manager and the management controllers located on the FRUs within a shelf. In an ideal world this would mean that an FRU manufactured by one vendor will interoperate with a shelf manager provided by another vendor, but to achieve this requires a significant integration effort because portions of the specification are interpreted differently resulting in incompatible implementations. A comprehensive compliance test suite that validates conformance to these specifications would go a long way to achieve interoperability.

There is also a need to define the interface between shelf manager and the system manager to achieve interoperability at that layer.

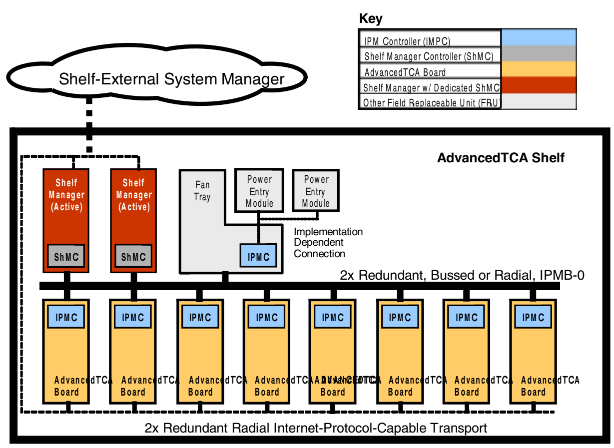

Figure 2: AdvancedTCA management architecture

Thermal Interoperability of Shelf and Boards

An AdvancedTCA integrator has to select a shelf that is capable of cooling a given configuration of boards under a broad range of operating conditions such as those defined in Network Equipment Building System (NEBS) deployment practice. The shelf suppliers rate the cooling capacity of their shelves, expressed in volumetric airflow through each slot. However, without an industry-wide benchmark, the underlying airflow tests and simulations are constructed differently from one vendor to another, and changing slightly just one parameter of the test, such as the slot thermal impedance, will significantly influence the reported ratings. This makes objective side-by-side comparison of shelves difficult, adding an element of risk to the interoperability of AdvancedTCA products in the ecosystem.

Adding to the challenge, the shelf airflow will be re-distributed under load (i.e., with boards installed) due to the fact that blade thermal impedance varies with the type of blade. Slots populated with low-impedance boards (e.g., switch blades) adjacent to high-density Single Board Computers (SBCs) will effectively starve the neighboring high-density slot of needed airflow, potentially exposing the SBC to thermal failures. Although PICMG3.0 mandates that board and shelf vendors publish pressure-airflow curves, it does not mandate impedance thresholds for boards.

Without a benchmark for consistent airflow rating of shelves, and without complementary guidelines mandating blade impedance thresholds, integrating a thermally sound platform remains an art of simulating and optimizing the P-Q curve of each slot to match the airflow requirements of its intended board. The industry desires a broader configuration-independent thermal interoperability model of shelves and boards.

THE INTEROPERABILITY FRAMEWORK

Interoperability Requirements –The Blueprint for Compliance

Interoperability requirements provide the link between options promoted by the PICMG specifications and interoperable products. The same holds true for the Service Availability Forum (SAF) and other specifications used to develop Modular Communications Platforms. The clear objective is to fill this gap so that vendors will deploy products that interoperate with each other, rather than customize SKUs for each design win. This linkage requires a narrowed view of the parent specifications to support a set of usage models. There are valid and important uses of the variety of options that PICMG promotes; however, some are applicable only to a small set of applications.

Interoperability of products relies foremost on conformance to standards. The interoperability requirements define the interface requirements from relevant standards organizations to set the stage for interoperability. As necessary, the performance levels required to support the target applications may also be defined. Beyond this level of specificity, interoperability requires one more step–conformance.

Conformance is assured when a product is validated to meet specific test criteria designed to ensure that the interoperability requirements are met. To reduce the overall cost of the compliance testing, it is important not to test beyond that which is required to demonstrate that the interoperability goals are met. To that end, test requirements can be classified into four categories, each level elevating the difficulty, hence cost, of the implementation.

1. *Inspection*–Requires physical examination of a Device Under Test (DUT) by a human operator to verify compliance.
2. *Connectivity*–External equipment is required to verify point-to-point connectivity of a DUT in a passive state.
3. *Functional*–Requires a DUT to be put into an operational state to verify compliance.
4. *Performance*–Higher degree of functional validation is required, such as signal integrity or stress.

These test objectives lead to a set of interrelated test procedures that ensure adequate validation to promote interoperability.

Interoperable Management

The PICMG 3.0 AdvancedTCA base specification defines the management infrastructure inside the AdvancedTCA Shelf in great detail. The PICMG AMC.0 Base specification further defines the infrastructure for managing the mezzanine modules. Ideally, the support of these requirements by respective components should produce interoperable systems, but that is not always the outcome. In some instances, the requirements are implemented differently leading to interoperability issues. In other cases, requirements do not exist. Therefore, a focused effort is necessary to achieve interoperability across components.

The first step is to list all the existing requirements in the PICMG specifications that are important for interoperability, and target them for compliance testing by a common test suite. A collaborative approach among component vendors drives consensus toward a common interpretation of the requirements, improving the likelihood that different vendors' implementations of the same requirement will conform, thereby improving interoperability.

Additionally, there are many areas which are of practical importance, but have not yet been addressed by the present specifications. To address these gaps identified by the system integrators and TEMs, new requirements have to be formulated.

One simple example is obtaining the revision number of various components on the boards such as hardware, firmware including the BIOS, FPGA etc. Currently there is no uniform way to query this information from the boards. Consequently, in some cases it may be necessary to reboot a board just to find its BIOS version information. To alleviate this situation, a new IPMI command can be defined for accessing the version and revision numbers of these components.

Another example is handling of telco alarms. Telco alarms are important for telecom systems. The current specification articulates the hardware-related requirements for support of telco alarms such as connector pin assignments and electrical specifications, but it does not provide a uniform way to control the generation and suppression of the alarms themselves. Here again, a new command and message structure needs to be defined to facilitate control of the telco alarms. This command can then be adopted by the system management entity to exercise control over the telco alarms.

There is a strong industry incentive to identify the problem areas for interoperability and agree on solutions to these challenges that translate into new requirements. Ultimately, these requirements will guide conformant implementations with enhanced interoperability.

The interface between shelf manager and system manager is also a major interoperability concern. The PICMG specification mandates the support of Remote Management Control Protocol (RMCP) at this interface. RMCP is Intelligent Platform Management Interface (IPMI) messaging over the User Datagram Protocol (UDP). However, a typical IPMI message is 32 bytes long, which is not adequate for the serious amount of communication required over this interface. Thus, one approach is to define another interface and leverage it to integrate the shelf manager with a system manager in an interoperable fashion. This scenario envisions expanding the current scope of PICMG and mandating a standard interface at this management layer, such as the Hardware Platform Interface (HPI) already defined by the Service Availability Forum (SAF). This brings about an interoperable interface architecture between the system manager and ATCA shelf manager, as represented in Figure 3.

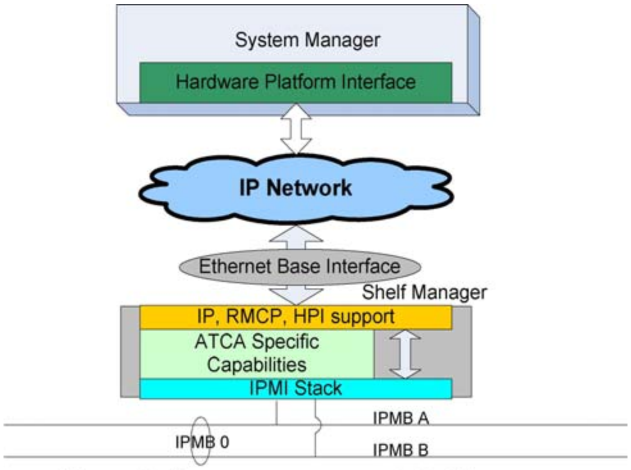

Figure 3: System manager and shelf manager interfaces

Thermals: Boards Meet Shelf

The PICMG3.0 specification recommends the upper thermal envelope as 200 Watts per slot. The laws of thermodynamics then dictate the minimum required volumetric flow rate sufficient to dissipate this thermal power, under given environmental conditions. Since board thermal impedance affects the air throughput within the slot, PICMG3.0 mandates that board and shelf vendors publish pressure-airflow curves, to aid in the shelf/board integration process. These P-Q curves are helpful in simulating the effects of low-pressure and high-pressure boards in various shelf slot configurations, allowing the shelf slot configuration to be customized for a particular set of boards in a given shelf. However, any board reconfigurations within the shelf will inevitably impact slot impedances, thus re-distributing the airflow across the slots. This falls short of enabling a generic reconfigurable shelf, without impacting the thermal integrity of the platform. The reason is simple–the boards are not required to use the available airflow judiciously, and shelves have no requirement to deliver an adequate airflow proportional to slot impedance.

To enable thermal interoperability across a broad cross-section of boards from different vendors, there is a need to mandate that boards present minimum impedance to airflow based on their power consumption. When implemented, this model allows for a far greater degree of thermal interoperability by providing balanced impedances across the shelf. This also facilitates moving disparate boards to different slots of the same shelf, to build targeted network applications, while preserving the thermal integrity of the system.

On the shelf side, the current PICMG3.0 recommendation of a 200 W/slot alone does not provide an adequate design requirement to guide the shelf vendor. Rather, the shelf's cooling capacity is determined by the amount of airflow delivered in each slot of the shelf. Therefore, there is a need to outline shelf thermal requirements explicitly in terms of the minimum airflow per slot that a shelf has to supply, pegged to a reference slot impedance. This philosophy could then be applied to measure slot airflow under the full range of environmental conditions as mandated by NEBS and the European Telecommunications Standards Institute (ETSI) specifications.

Laying out this interlocking framework for blades and shelves will enable a deterministic ecosystem of shelves and boards designed for thermal interoperability from the ground up.

INTEROPERABILITY FRAMEWORK

Test Specification & Tools

The identification of interoperability requirements is the first step towards achieving interoperability in the multi-vendor AdvancedTCA environment. The next step is compliance testing for these requirements. This is a significant task, keeping in mind the scope of AdvancedTCA architecture, which consists of mechanical, thermal, electrical and manageability components. Each of these areas requires its own kind of test environment.

The compliance test requires very precise, verifiable, and repeatable test procedures. These procedures need to be developed in a collaborative environment where different stakeholders such as component manufacturers, system integrators and TEMs bring their own insight to the problem. The requirements specified in the first step will lead to the test procedures that will be organized into a test manual. The test manual defines the environment, tools, and detailed steps to validate these requirements.

Automated Test Suite

Manageability requirements describe software interaction and as such can be validated via automated software test suites. The AdvancedTCA manageability framework is based upon IPMI infrastructure. All the intelligent FRUs are linked to the shelf manager by IPMB links and the shelf manager exchanges IPMI messages with these FRUs. The software test suite exchanges IPMI messages with the shelf manager as well as management controllers on the intelligent FRUs to exercise this interface. A conceptual view of IPMI test architecture is provided in Figure 4.

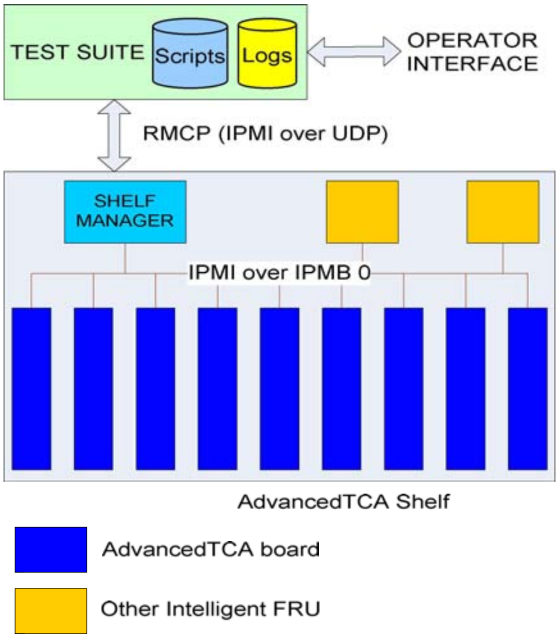

Figure 4: IPMI test architecture

In keeping with the spirit of open standards, this software test suite must be deemed impartial and be provided by a neutral vendor, one that does not own a stake in certification of their own AdvancedTCA products. There are already some test tools available in the industry which can be leveraged to develop an interoperability compliance test suite.

Intel offers its IPMI V2.0 Conformance Test Suite (ICTS), available for download in the public domain: http://www.intel.com/design/servers/ipmi/tools.htm. The ICTS was conceived for testing of basic IPMI compliance, and the IPMI messaging infrastructure provided by this suite can be leveraged to build other IPMI-based test tools. Another suite based on ICTS is called AdvancedTCA Compliance Test Suite (ACTS), a test framework focused on compliance test for a subset of requirements from the PICMG3.0 base specification.

There are other industry-recognized test suites that target requirements from the PICMG 3.0 specification. They represent a significant body of work that can be built upon when formulating a systematic interoperability compliance.

The test suite for interoperability will span beyond the PICMG defined domain; therefore, the test suite will include other types of tests in addition to IPMI-based interaction. Two such prominent areas from user standpoint are suites for HPI and SNMP as standard interfaces for the shelf manager. Architecturally, the IPMI, HPI and SNMP domains should be based on a common framework, unified by a user interface and common logging facility, referenced in Figure 5. This extensible test framework will evolve in modular fashion to keep up with future needs.

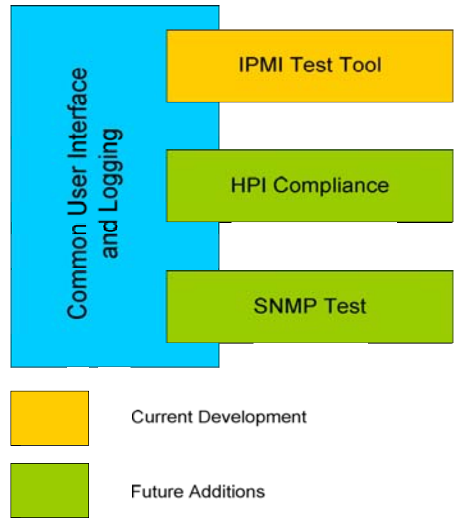

Figure 5: Automated test suite

Other Types of Tools

Other areas of interoperability testing don't lend themselves to automated software tools, rather they require golden test platforms to be defined and instrumented. One such application is the interoperability requirements framework of thermal requirements for AdvancedTCA shelves. These requirements share the principle that accurate slot airflow analysis is representative of a shelf's cooling capacity, and thus can be validated using a common airflow test methodology. In consultations with the industry, Intel has already implemented a functional test platform of this type.

The purpose of the flow benchmark is to measure the available airflow in each slot of the shelf, under a specified impedance load. The test platform consists of the shelf under test, provisioned with a standard reference board serving several functions:

- It provides for laminar airflow within the slot.
- It models the P-Q characteristics of a typical SBC.
- It facilitates CFM measurements across the depth of the board.

All slots of the shelf under test are configured with the standard reference board, presenting a pressure drop of 0.15 inches of water. The key advantage of this methodology is that it employs a managed and fully functional shelf, thus taking empirical flow measurements under real-life conditions, Examining Figure 6, it can be seen that several air pressure drops develop as the airflow travels from the shelf inlet to the shelf outlet.

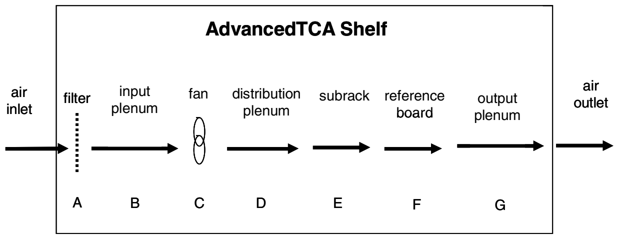

Figure 6: Airflow pressure drop reference diagram

These drops exist regardless of the location of the fans or blowers or any other component of the shelf. Measuring the volumetric airflow directly in each slot equipped with the reference board provides a very accurate picture of the actual amount of air that reaches the slot. This represents the airflow available for cooling each board of the shelf.

An example of the resulting shelf airflow profile for a 14-slot shelf is shown in Figure 7.

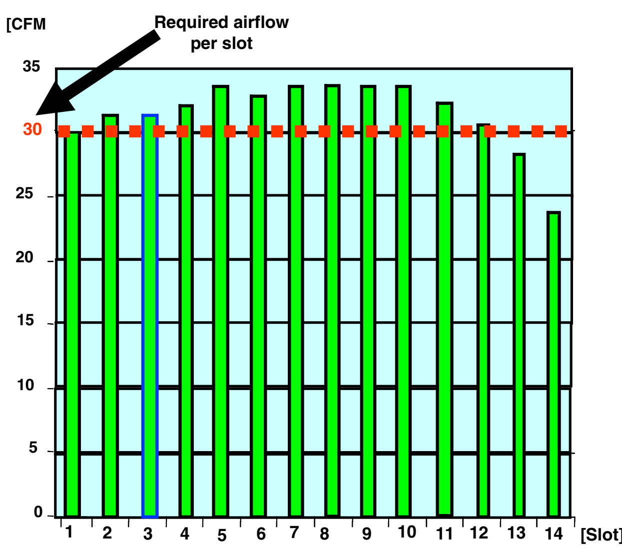

Figure 7: A typical shelf airflow profile

Mandating airflow requirements and adopting industry benchmarks to validate compliance will provide much desired transparency into the thermal capabilities of the shelf ecosystem, well in advance of the actual platform integration.

Another category of tools will also be required for electrical testing of backplane and interconnects. A small subset of environmental testing will likely be integral to the interoperability conformance validation. For instance, an acoustic chamber may be required for sound power and sound pressure testing. This is to ensure that a shelf meets regulatory requirements, while maintaining its capability to deliver sufficient airflow per slot.

Certification of Interoperability Compliance

Certifying for interoperability is a concept already established in many other technology areas. The AdvancedTCA industry has yet to organize itself around a functional certification program.

In addition to identifying interoperability requirements, and the creation of detailed test procedures and underlying test tools, there are two other components to an interoperability program. The first is to coalesce on a certification process, yet to be determined. The second challenge is to set up an operational certification infrastructure with all the necessary resources and specialized test equipment such as described in this paper.

Examples of well-executed interoperability certification programs are found across many industries. Much can be learned from the well-established certification program and labs for Ethernet interoperability testing, which are closely akin to the component certification expected for AdvancedTCA. Yet the unique demands for an AdvancedTCA interoperability certification will come from its expected breadth and relative complexity. Since the AdvancedTCA architecture encompasses electrical, mechanical, thermal, interconnect, and software areas, each type of component has to be tested for these attributes. This calls for a robust lab infrastructure, with critical mass of specialized test equipment, as well as personnel that are versed in all aspects of AdvancedTCA testing.

Founding of an interoperability certification program is the next logical step towards achieving the goal of multi-vendor interoperability in the AdvancedTCA ecosystem. Together, the industry will be able to fund the necessary infrastructure, as well as contribute its collective expertise to this noble cause.

CONCLUSION

As growing numbers of integrators race to deliver solutions on AdvancedTCA platforms, focusing on interoperability of components from different vendors becomes paramount. While there have been early efforts at standard compliance and sporadic interoperability testing, they have not been adequate to fully validate, document, or certify the level of interoperability deemed necessary by the integrators, and ultimately their customers, to help them launch AdvancedTCA platforms faster.

To mature to the next level, the AdvancedTCA industry must bring together the component suppliers, system vendors, and service providers in a concerted effort to specify, develop, certify, and deploy interoperable modular solutions. The identification of interoperability requirements takes the first step toward an industry-wide interoperability program. Born from broad industry collaboration, the first set of such requirements comprehends the basic interoperability requirements relevant across all the AdvancedTCA implementations, independent of any specific application segment. Later on, profile-specific interoperability requirements will be added for different segments. As shown, the supporting infrastructure behind a certification program must also include test procedures linked into the requirement set, test tools, and benchmarks as well as an operational interoperability lab, all governed by efficient certification processes.

As AdvancedTCA components designed and certified for baseline interoperability become available, integrators will have the flexibility to cost-effectively mix and match modular components from different vendors. The ultimate success of an interoperability certification program will be measured on the economic benefits and agility that a thriving interoperable ecosystem will bring to service providers.

ACKNOWLEDGMENTS

We acknowledge all members of the Modular Communication Platform Division who share the vision of AdvancedTCA-based interoperable Modular Communications Platforms.

REFERENCES

[1] PICMG 3.0 "AdvancedTCA Specification 2.0," March 2005.

[2] "InfiniBand Architecture, Volume 1. General Specifications," *InfiniBand Trade Association*, 2000.

[3] "InfiniBand Architecture, Volume 2b. Mechanical, Cable and Optical Specifications," *InfiniBand Trade Association*, 2000.

[4] PICMG 3.1 "Specification for Ethernet fabric Interconnects," December 2002.

[5] PICMG 3.2 "Specification for InfiniBand fabric interconnects," January 2003.

[6] PICMG 3.3 "Specification for StarFabric interconnects" January 2003.

[7] PICMG 3.4 "Specification for PCI Express interconnects," January 2003.

[8] Service Availability* Forum "Hardware Platform Interface," March 2004.

[9] "The Evolution of WiMAX Certification," *Monica Paolini, Senza Fili Consulting*, October 2005.

GLOSSARY

Interoperability	The ability of two components to provide intended functionality between the components while operating in a common system environment. The ability of software and hardware on different, but interfaced, machines to communicate with each other.
LVDS	Low voltage differential signaling.
M-LVDS	A later development of LVDS defined in TIA/EIA-644. It is specifically designed for multisource and multidrop signaling.
Shelf	AdvancedTCA shelf consists of subrack, backplane, front boards, cooling devices, rear transition modules, power supplies, etc. Also historically known as chassis.
System	A collection of components organized to accomplish a specific function or set of functions. For the purpose of this discussion it is a configuration of AdvancedTCA components such as shelf assembly, shelf manager, front boards and/or AMCs in order to achieve a specific function such as telco server.

AUTHORS' BIOGRAPHIES

Steve DeNies is a principal engineer in the Communications Infrastructure Group of DEG. His expertise in boards and systems spans multiple levels including architecture, design, and launch. He has focused on AdvancedTCA architecture since 2002. He holds a BSEE degree from the State University of New York at Buffalo. His e-mail is steve.denies at intel.com.

Jay Gilbert is a senior technical marketing manager within Intel's Communications Infrastructure Group. He is responsible for standards development, technology evangelism, and ecosystem enabling in support of Intel's AdvancedTCA communications infrastructure development efforts. Jay has a BSEE degree from the

Oregon Institute of Technology. He also holds an MBA from Portland State University. He can be reached at jay.gilbert at intel.com.

Ashok Mishra is a system architect in the Communications Infrastructure Group of DEG. His current interests include AdvancedTCA-based next-generation telecom platforms and their management architecture. He holds a BSEE degree from the National Institute of Technology, Allahabad, India. His e-mail is ashok.mishra at intel.com.

Karel Rasovsky is a technical marketing manager in the Communications Infrastructure Group of DEG, involved in industry evangelism of modular communication platforms. His prior experience at major TEMs spans network infrastructure planning and launch, product line management, initiative marketing, strategic business planning, and competitive intelligence. He holds an MS degree in Computer Engineering from Florida Atlantic University, Florida, and a BSEE degree from the University of Technology in Brno, Czech Republic. His e-mail is karel.rasovsky at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

Quality Campus VoIP: An Intel® Case Study

Ranjan Sinha, Information Technology, Intel Corporation
Catherine Spence, Information Technology, Intel Corporation
Tim Verrall, Information Technology, Intel Corporation

Index words: VoIP, Campus Voice, QoS, SIP, Host Media Processing, PIMG

ABSTRACT

As IT departments and the communications industry move to a single converged voice and data network solution, the demand on the network to provide specific service levels increases. This demand for higher and differentiated service levels for certain applications has created a need for a reliable and scalable Quality of Service (QoS) model in the network. Prior to the introduction of Voice over IP (VoIP) into a production network, the environment must be equipped to support the reliable transport of voice and a solid end-to-end voice quality plan must be defined and in place.

Typical demands on the network when supporting voice are continuous uptime (99.999%) and consistently low deviations in latency (jitter). As a real-time application, VoIP is far more sensitive to latency, jitter, and packet loss than standard data applications as users have no tolerance for garbled or broken speech. VoIP continues to be compared to traditional telephony in terms of voice quality and reliability as this is the level of service quality to which most users are accustomed.

QoS is a descriptive for defining how IP packets are dealt with through network devices. Mean Opinion Score (MOS) is used to measure the "quality" of a telephone call. QoS on the network helps facilitate a better MOS score for voice on an IP network. QoS is not a standard or protocol, but simply a generic industry term for outlining technologies, standards, and strategies to provide for network quality. In general QoS, to facilitate good voice quality and high MOS, requires that packets carrying real-time voice traffic cannot be delayed and must be prioritized over data traffic, which can better tolerate being slightly delayed. Most jitter in the network is caused by queuing delays associated with momentary or chronic congestion. QoS for voice can help make this queuing delay transparent to the voice packets.

This paper examines a case study based on Intel's experience of deploying VoIP [2] with voice quality within a campus and converging voice and data on the existing local area network (LAN). The methodology used to enhance end-to-end voice quality includes ensuring bandwidth, enabling QoS and optimizing IP telephony endpoints. This case study demonstrates that even a basic voice quality and QoS strategy can produce high-quality results with minimal infrastructure upgrades.

INTRODUCTION

In this paper we examine a case study based on Intel's experience of deploying Voice over Internet Protocol (VoIP) with quality voice within a campus where voice and data were converged on the existing LAN using Session Initiation Protocol (SIP). We demonstrate the voice quality methodology used and the benefit it provided. First, we familiarize the reader with the terminology used throughout the paper and provide an overview of the Intel case study. Next, we discuss the network infrastructure elements, voice quality plan specifics, and trial results. Finally, we explore the challenges and solutions and summarize our key learnings.

Background and Terminology

SIP is an Internet Engineering Task Force (IETF) protocol that is used to initiate interactive user sessions with multimedia elements. SIP is specified in IETF Request for Comments [RFC] 2543 [1]. In the Open Systems Interconnection (OSI) communications model, SIP takes place in the Application Layer (Layer 7) and is responsible for establishing, modifying, and terminating the user sessions. In this case study, the user sessions are Internet telephony phone calls.

In a traditional telephone system for a campus, a PBX or Private Branch Exchange is used to provide call switching that is circuit-switched. Functionality typically provided by traditional PBX systems includes local and least-cost call routing, call forwarding, low-density call conferencing, and call detail recording. Most traditional PBX systems use a proprietary digital protocol over a private local area telecommunications network. Cables are run from the PBX to telephone stations in offices and cubicles, often in parallel to data network cables. Voice quality in a traditional voice network is usually a non issue because the end-to-end path is ensured for the duration of the phone call. The dedicated network provides low latency. Jitter is very low since the paths of the signals take the same route throughout the system. Physical faults such as loose connections or cable malfunctions can degrade the quality of the call. However, these impairments are detected by monitoring systems that measure electrical characteristics of the system.

In a VoIP solution for a campus, an IP PBX is used to provide the local telephone system. Because of the greater access to data and the incorporation of open standards, IP PBX systems generally provide the same features as traditional systems with more intelligence and there is greater opportunity to integrate with standard business applications enabling a higher level of automation. Sometimes media processing applications such as voicemail and automated attendant and Interactive Voice Response (IVR) are converged on the IP PBX, making adjunct devices unnecessary. Physically, the IP PBX may be connected to the data LAN or to a private voice LAN. Monitoring call quality for the IP PBX system becomes more important than in a traditional PBX, due to the nature of packetized voice and variability introduced by traditional networking.

MOS is used to measure the *quality* of a telephone call, whether it is a traditional circuit-switched or IP telephony call. MOS is determined by a panel of human listeners in a controlled environment who rate the audio from 1 to 5, with 5 indicating best quality. A MOS of 4 is considered acceptable quality, where 90% of the users are satisfied with the quality of the call.

QoS refers to how IP packets are dealt with through network devices. QoS on the network helps facilitate a better MOS for voice on an IP network. QoS is not a standard or protocol, but simply a generic industry term for outlining technologies, standards, and strategies to provide for network quality. In general, QoS, to facilitate good voice quality and high MOS, requires that packets carrying real-time voice traffic cannot be delayed and must be prioritized over data traffic, which can better tolerate being slightly delayed. Most jitter in the network is caused by queuing delays associated with momentary or chronic congestion. QoS for voice can help make this queuing delay transparent to the voice packets.

Introducing the Intel® VoIP Program

The Intel® VoIP program took place at one Intel site over a six-month period, with over 50 participants who represented a variety of job functions and telephone usage models. The VoIP system was deployed within the enterprise with standard business voice applications, including automated attendant, voice mail, unified messaging, follow-me (call forwarding) features, fax, and remote access.

The IP PBX was built on Intel architecture with additional communication building blocks from Intel. The range of telephone devices used in this trial enabled the demonstration of SIP interoperability of IP endpoints. Additionally, the IP PBX was integrated with Intel's standard Instant Messaging (IM) solution to share basic user presence, along with the ability to launch a phone call.

The project met each of several high-level business goals that were set. Those goals included using and validating some of Intel's VoIP products, giving Intel's IT department the opportunity to introduce a converged communication solution into the enterprise, and demonstrating a phased migration path from a traditional telephone system to a next-generation IP PBX. How the goals were accomplished is outlined below.

- Validate Intel products.

  Through the success of the project, Intel was able to validate that an open standards VoIP solution using SIP can be deployed with Intel products. Intel is maintaining a showcase platform and demo room at the pilot site where interested parties from inside and outside Intel can experience the converged communication solution firsthand.

  The Intel components used in the solution are as follows:

  - The Intel® Xeon® processor is incorporated in the industry-standard server running the IP PBX
  - Intel NetStructure® Host Media Processing (HMP) Software within the IP PBX performs media processing tasks on Intel servers without the use of specialized hardware.
  - Intel NetStructure PBX-IP Media Gateway (PIMG) is an Internet appliance that integrates

® Intel, Xeon, and Intel NetStructure are trademarks or registered trademarks of Intel Corporation of its subsidiaries in the United States and other countries.

the proprietary digital telephones with the IP PBX through telephone set emulation.

- Intel® Centrino® mobile technology on laptops runs the IP PBX client software and softphones.

- Introduce converged communications.

  Intel introduced Intel technology into the enterprise for converged communications at three levels. First, applications were converged at the server including base telephone system and media processing applications. Second, voice and data were converged onto the same LAN. Finally, telephone applications were converged at the client including call control and softphone.

- Demonstrate phased migration.

  The production showcase environment demonstrates to customers and others outside Intel how a phased migration from a legacy PBX to IP PBX might proceed. In this case, the IP PBX was deployed behind the traditional PBX. Voice was layered on top of the existing data LAN.

Planning for and monitoring voice quality were critical elements of the trial. Because users are accustomed to the high-quality performance of a traditional telephone system, they have no tolerance for echo, jitter, or delay in VoIP communication. As a real-time application, VoIP requires the same continuous uptime as the system it replaces. High quality is possible, provided careful planning goes into the network design and implementation details.

VOICE INFRASTRUCTURE

In this section we provide an overview of the technical voice infrastructure with an emphasis on characteristics important for voice quality and QoS, which include connecting the IP PBX with the PSTN, gateway integration, and a general discussion of the IP telephony endpoints.

The IP PBX resides behind the traditional PBX, meaning that no changes were needed to provide the connection between the premises and the public telephone network. All calls continue to come into and out of the campus through the existing PBX. The PBX continues to route inbound calls as usual, directly or through the auto-attendant without modification. This configuration enabled quick installation of the new system and removal of the old, with minimal disruption.

To make the connection between the traditional PBX and the IP PBX, a gateway is used to enable translation of TDM circuit-switched calls from the existing PBX to packet-switched calls from the new IP PBX. In this case the direct physical connection to the PBX is through a PRI ISDN T1 interface. The gateway is connected to the LAN through a switch, as shown in Figure 1, making sure that the configured protocols are matched up properly.

Pilot users tested a variety of IP telephony endpoints, including existing legacy phones and IP phones. Most users received IP phones. Legacy phones are connected to the IP PBX using Intel's PIMG as shown in Figure 2. IP phones are connected directly to the LAN.

® Intel and Centrino are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

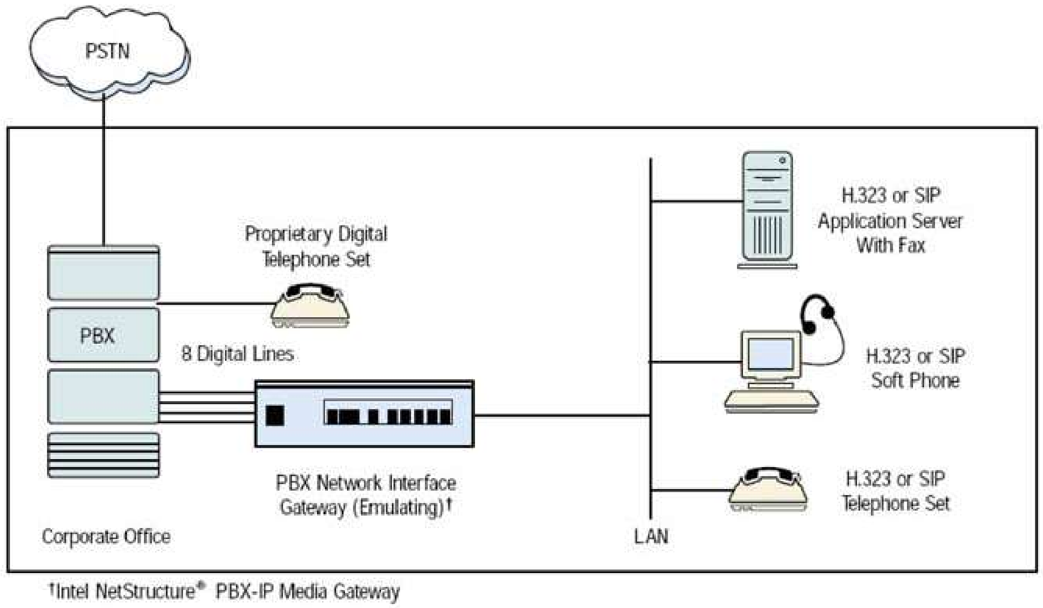

Figure 1: Connecting the IP PBX to the traditional PBX

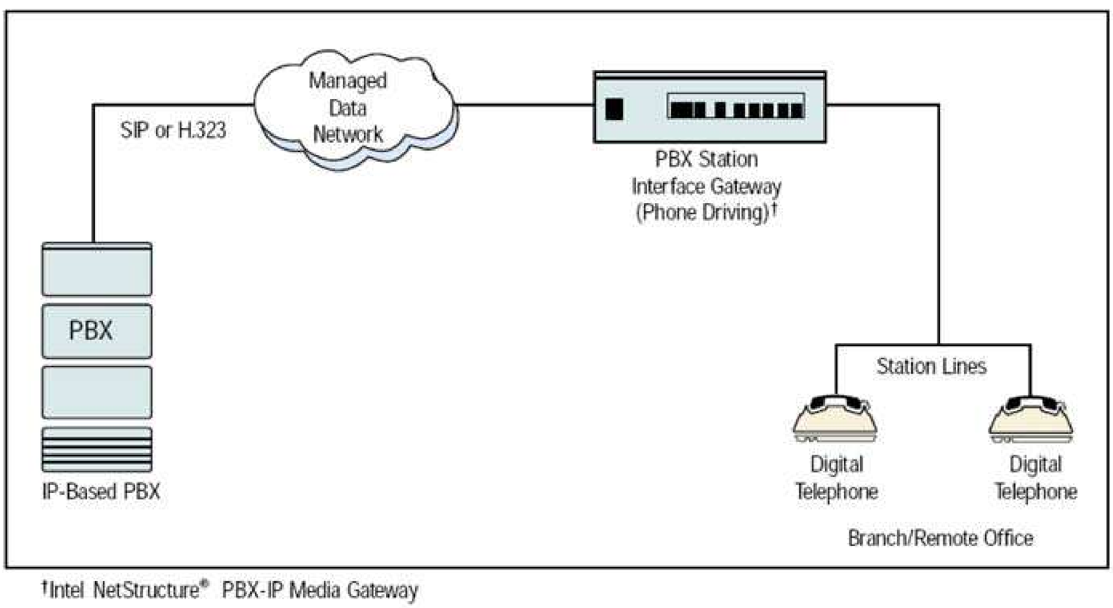

Figure 2: Using PIMG to connect legacy phones

We tested several types of IP-based phones in the pilot, including hardware phones, software phones, and wireless phones. All these were SIP-enabled out of the box. Of users with IP phones, 58% received a hardphone, 28% received a softphone, and 14% received both. Wireless IP phones were tested but not deployed to any live users.

Some hardware IP phones were deployed with Power over Ethernet (PoE). PoE makes for a more streamlined approach since a separate AC adapter for each phone is expensive and takes up valuable outlet space. PoE enables the transmission of electrical power over the same Ethernet cable as voice and data, using industry-standard IEEE 802.3af. For implementation, the access switch had to support the PoE feature, so in some instances the switch was replaced in this trial. An advantage of PoE is that it offers backup power capabilities if power goes out in the facility. Since the PoE cable draws power from the network, power during an outage continues to be supplied via the network's Uninterruptible Power Supply (UPS). This improves the user experience and overall perception of voice quality.

Another important feature of the hardware IP phones with respect to QoS is the ability to do 802.1q virtual LAN (VLAN) trunking and the ability to mark outgoing packets using differentiated services or *DiffServ*. In areas where these phones were on a separate VLAN, the default VLAN was used. In others, 802.1q and DiffServ were used.

Software phones or softphones consist of a software client installed on a PC, connected to a headset with audio and microphone. Good quality audio devices are essential for a good end user experience. The softphones, like hardware IP phones, have the ability to mark the packet with DiffServ allowing those packets to obtain priority on a properly configured network. However, there were other factors that also impacted voice quality. For example, Intel tested wired and wireless headsets from a variety of vendors and found that the selection had a profound impact on call quality and usability.

- **Wireless headsets**–Several types of Bluetooth[*] wireless headsets were tested and none were able to provide acceptable call quality. Noise and echo were problematic. Further, there were major usability issues concerning power including battery life and awkwardness of turning power on to answer an incoming call.
- **Wired headsets**–Wired headsets come either with their own Digital Signal Processor (DSP) using the PC's USB port, or they use the PC soundcard through the headset and microphone plugs. Users experienced a wide range of quality issues including noise and echo. The more expensive USB headsets with DSP provided the best quality.

VOICE QUALITY PLAN

The plan was very simple and straightforward. The goal was to provide QoS for the voice traffic end-to-end across the IP LAN via a DiffServ model, using Class of Service (CoS) and Differentiated Services Code Point (DSCP) where appropriate. DSCP refers to how the network traffic is marked or *tagged* in the IP layer (Layer 3) to identify it as part of a particular class. Voice packets are marked with DSCP (DiffServ code point) on the end device; this marking is mapped to a CoS tag in the LAN switch. CoS refers to categories of network traffic that are recognized and queued by the network switches. CoS marking or tagging occurs in the Layer 2 header for the 802.1q trunks, and this method is referred to as 802.1p DiffServ, as the overall model defines how these packets are queued and forwarded in the network. Based on the policies set, network switches use CoS and DSCP values to provide high-priority handling for voice traffic.

The life of an audio packet as it is transported over an IP network is given here:

- Human speaks
- Coded with some form of compression
- Inserted into packets with sequence numbers
- Transported onto the network
- Received into network equipment
- Network determines next-hop interface and switches the packet
- Packet is queued for sending to the next hop
- Received in a playout buffer
- Decoded in sequential order
- Played back to human

As packets traverse the network, all elements must recognize which packets are real-time audio and prioritize their handling, yet ensure that no uplink, or any other portion of the network, becomes over utilized by the addition of voice on the network.

Another factor is the coder-decoder or *codec* used to compress and decompress voice transmission. G.711 is a standard for speech codecs that provides toll quality audio at 64 Kbps. Intel elected to use uncompressed G.711 as a codec in this trial since it provides the highest quality available for the system implemented, and bandwidth utilization is not an issue on the LAN.

LAN DESIGN

In this section we provide the architectural details on how the LAN was organized and discuss how we supported the QoS plan. Our goal was to ensure that as little jitter or loss as possible was introduced into the network and that QoS was provided for all voice calls over as much of the network as was possible. Small amounts of jitter could be compensated for in the receive buffers. Small amounts of loss are handled by the codec, although this loss was very little.

The network was designed using Ethernet switches with dedicated full duplex ports for each endpoint device. The LAN design (Figure 3) consists of the core, distribution, and access switches. All the routing is done in the core; the distribution acts as the aggregation point for all the

[*] Bluetooth is a trademark owned by its proprietor and used by Intel Corporation under license.

access switches out in the communication closets, and the access layer is the connection point for all the users. In this particular design, we focused on four communication closets spread out across the campus. All connectivity between switches is Gigabit Ethernet linked by fiber. The access switches all deliver 100 Mb switched to the desktop and phones.

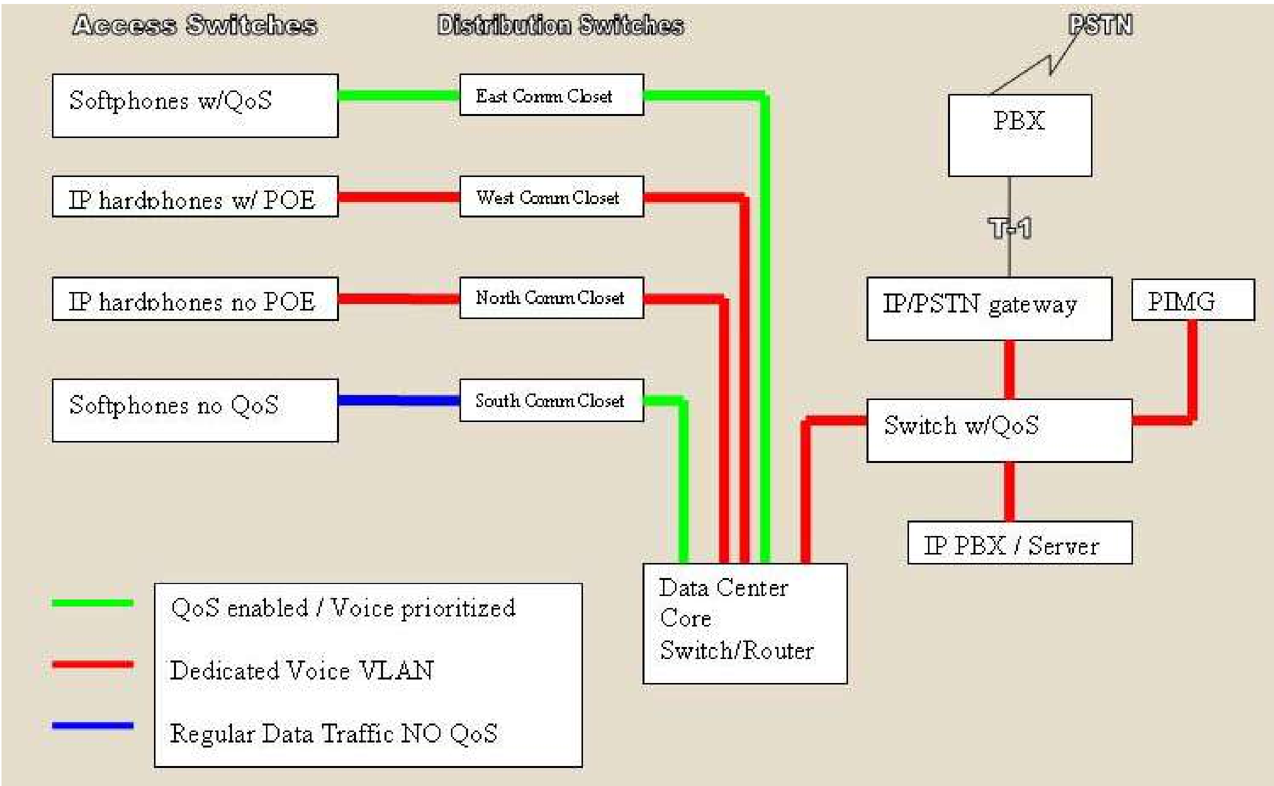

Figure 3: LAN design diagram

A switch located in the data center connects the IP PBX, the Gateway between the IP PBX, and the traditional PBX and the PIMG. All these devices are on the same VLAN, and QoS was enabled for any voice traffic in this segment. This allows for the end-to-end QoS for all the endpoints except for one communications closet.

In the network supporting the pilot user community there is a mixture of PoE and non-PoE access switches in the communication closets. There are four closets: North, South, East, and West, providing access for four different types of use cases identified:

1. Softphones with QoS: All the users in the East communications closet were using softphones on their laptops with headsets. For these users, it was essential to prioritize the voice traffic over their data traffic. This meant enabling QoS on the switches end-to-end from this closet. At the client, we installed the packet scheduler and made sure that the softphone client is marking packets.
2. Hardware IP Phones with PoE: All the users in the West communications closet were using hardware IP phones with PoE. The access switch in the West closet had to be capable of providing the industry-standard 802.3af PoE. The QoS mechanism used in this case was to segment all the hardware IP phones onto a dedicated VLAN using 802.1q and to use priority queuing for voice, to prevent contention with the data traffic.
3. Hardware IP phones without PoE: All the users in the North communications closet were using hardware IP phones but did not have PoE capability. These phones use AC adapters. The QoS mechanism used is the same as that used in the West communications closet including segmented VLANs and priority queuing.
4. Softphones without QoS: The South communications closet did not contain access switches that were capable of prioritizing voice traffic before forwarding it on to the distribution switches. Therefore, we deployed softphones without QoS to any users serviced out of the South closet. This model enabled us to contrast voice quality for users that had some form of QoS mechanism in place end-to-end versus those users that did not have end-to-end QoS.

In addition, we tested Wi-Fi phones in the lab with a dedicated access point so there was no other traffic but voice. The voice quality results were good but we did not perform extensive testing since we had a bigger problem to solve: secure device authentication was not ready in time to deploy users with phones. We decided instead to push Wi-Fi phone deployment to the next phase of the project.

Where the switch was capable of 802.1q and DiffServ, the PCs were also plugged into the switch of the hardware IP phones. This allowed the PC to be set up on a different VLAN than the hardware IP phone. This approach is excellent for a campus with wire runs of one port per cubicle/office or in cases where limiting the number of access switch ports is required to support both voice and data. With this approach the same functionality of using two different network ports is realized at half the real estate.

In summary, the philosophy behind the design was to ensure QoS on a link-by-link basis end-to-end. We enabled DiffServ, which enabled the prioritization of voice traffic over data traffic anywhere that there might have been contention for the network. VLAN separation was also enabled to provide easy differentiation and greater security of the voice network.

The uplinks were all Gigabit and lightly utilized, so we were able to ensure that enough bandwidth for the voice calls was available. Figure 4 shows the portion of the network with DiffServ enabled and how the prioritized packet travels through access, distribution, and the core.

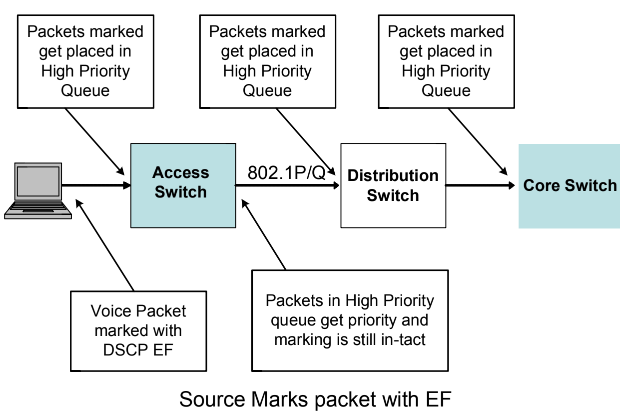

Figure 4: End-to-end packet flow

Figure 5 shows the relationship between the voice quality and various components.

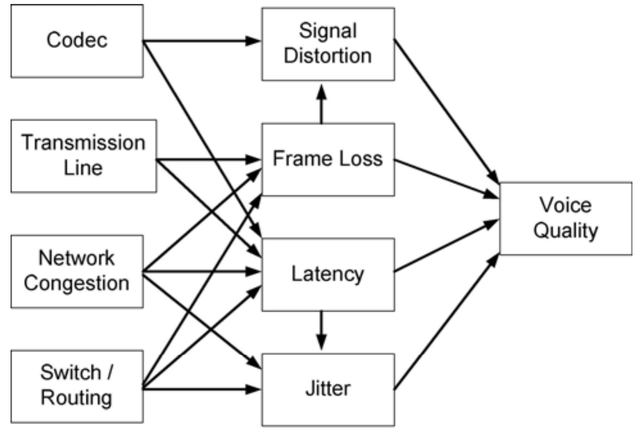

Figure 5: How voice quality maps to components

QoS must be comprehended by all IP telephony end points. All IP endpoints in the pilot were QoS-aware, including all IP phones, softphones, and Wi-Fi IP phones. The following sections describe the QoS considerations we followed for each type of endpoint.

Hardware IP Phones

We used a virtual VLAN segmentation (IEEE 802.1Q) to create a separate broadcast domain. That allowed us to segment voice traffic from network broadcast messages and data traffic. We used a separate VLAN for each supported network access switch. In other words, there was a VLAN for each communications closet.

Softphones

For softphones, we enabled end-to-end QoS in every component in the network. We set up the network and all of its components to distinguish voice and prioritize it above other kinds of traffic. The IEEE standards for QoS are 802.1p and 802.1Q, where a Layer 2 tag is used to specify packet priority. In 802.1p, CoS 5 identifies voice traffic. Using DiffServ, the voice packet is marked with Class EF, which gets priority on a QoS-enabled network. Unlike a hardphone, the softphone resides on a laptop or other computer that is probably running other applications. Like the network, the client environment must recognize and prioritize voice, and the selected softphone must support it.

Non-IP Hardware Phones

Part of the QoS plan included identifying and segmenting the voice traffic on the network from data traffic, and giving voice traffic its own dedicated, logical route. We performed that sort of segmentation only for IP hardphones, not for all voice traffic.

TRIAL VOICE QUALITY RESULTS

This section contains the call quality results experienced in this Intel case study. User reported results are discussed as well as management, reliability, and quality.

User Reported Results

Intel did not perform a formal MOS measurement with users. Feedback was collected through channels such as a weekly health check e-mail (yes/no voting buttons to indicate problems), a detailed qualitative user survey, and interviews conducted by a Human Factors Engineer.

Softphone with QoS users experienced good voice quality all the time, no matter what time of day whether in a conference call, outbound, or inbound call. These users were extremely satisfied with the quality of voice and in fact, with properly tuned audio settings on the laptops and good quality headsets, the voice quality was described as an improvement over their regular phones.

Voice traffic for hardware IP phone users had virtually no contention with the data traffic. These users also reported good quality voice. The make and model of hardware IP phones was important as we deployed a variety from multiple vendors. Users were dissatisfied with certain types of phones, so by trial and error Intel could determine which phones worked best in the environment. Once phones with reported issues were replaced, users described the quality of voice as very good at all times.

Where QoS capability between the access switches and the distribution switches was not available, we had some softphone users. These users expressed satisfaction with the quality of voice most of the time; however, they frequently experienced degraded quality of voice, reportedly several times per day. Often this was explained by the fact that heavy downloads were occurring on the access switches at the time of the calls. Without prioritization for voice traffic, call quality was impacted, even with the high availability of bandwidth in the LAN. This confirmed the necessity of enabling end-to-end QoS.

Users with legacy phones connected to the IP PBX through the PIMG also reported good voice quality. The only area that could be impacted through the voice quality plan was at the data center switch, where voice was prioritized and a separate VLAN was configured. The existing wiring was used between the PIMG and the legacy phones, ensuring dedicated connections.

Management, Reliability, and Quality

The reliability of the LAN was 99.98%. This was achieved through pro-active monitoring. There was no physical redundancy for the access switches. The traffic levels on the network were well understood through the baselining efforts and there were no bottlenecks identified. The utilization on the uplinks from the distribution to the core did not peak above 40% and the average utilization remained below 10%. Tools were in place to monitor the health of the network including visibility into the voice traffic that received priority. All management, monitoring, and alerting is centralized, providing end-to-end visibility into the network. From a management standpoint, the ability to push QoS parameters throughout the network from a central location is important in order to effectively manage a large deployment.

KEY CHALLENGES AND SOLUTIONS

Many challenges were found during this project, many were solved and some are as yet not completely resolved. We examine key challenges in this section including user perceptions of voice quality, softphone QoS issues, and voice quality impact resulting from security and reliability issues.

User Perceptions of Voice Quality

When requesting a user community to report all voice quality issues, you can expect many normally unreported instances to find their way into the report. This presents a challenge. The question is how to separate the real quality issues from the minor issues or those quality issues that have always existed. We did this by utilizing good proactive measurements of the network, evaluating the QoS usage and load along with making regular test calls and qualitatively measuring the quality through user feedback.

Softphone Quality Issues

This is an area where a number of challenges exist and are not all solved at this time. Some of these issues relate more to the quality the user perceives for the voice call than anything technical. The application gaining higher priority of the PC resources is one aspect that still needs to be solved, although many CPU management tools exist. The system should provide QoS for VoIP traffic out of the LAN interface card on the PC itself. The QoS packet scheduler on the PC is a step in the right direction, but a user can place all his/her client traffic in the voice (expedited forwarding) queue and so a centrally managed system is required before the client QoS can really be trusted. There were many areas where we found attractive and simple solutions. For example, it was easy to replace headsets where quality and user acceptance was much better when we used headsets with DSP.

Security and Reliability Issues and QoS Impact

There can be many potential security issues relating to a VoIP design and deployment and some of these played a role in this case study. First there was the question regarding what happens to VoIP if there is a worm or a Denial of Service (DoS) attack on the network. This problem was significantly reduced by providing a dedicated QoS protected VLAN for all the hardware IP phones. This keeps the data traffic separate from the voice traffic and therefore would allow the voice traffic guaranteed access to the network during a DoS attack.

Another area of IT concern is privacy. IP phone calls are much easier to tap and listen in on. If end-to-end encryption is used to protect the content then it makes it more difficult for the network to distinguish and prioritize voice. Also, encryption can introduce latency. One possible solution is to not converge voice on the LAN and instead install a private voice network. However, that approach significantly degrades the business value of converged networks.

CONCLUSION

The Intel VoIP program was a very successful trial on how to introduce VoIP into the enterprise using SIP and converging voice and data on the LAN. As the telephony industry moves from vertically integrated proprietary solutions to horizontal open standards architecture, a key challenge is to provide quality and reliable voice service on a mixed IP network. This project implemented a very basic voice quality and QoS strategy, proving that VoIP can be layered on top of an existing infrastructure with minor upgrades.

In terms of voice quality and QoS, the key learnings from this trial were as follows:

Voice Quality Plan: Intel benefited from having a deliberate plan up front. Enabling QoS leads to better and more predictable voice quality. This is a mandatory item for future deployments of any VoIP implementation in Intel. It is important to address voice quality holistically, from voice endpoint to voice endpoint.

Loading VoIP on a production LAN: Adding voice to an existing data LAN was important because it showed that the voice traffic did not impact the LAN functionality. It enabled the extrapolation of how scaling the pilot impacts the LAN from a capacity planning perspective. In addition, Intel received the benefit of managing only one network–one of the most touted benefits of VoIP. One major issue yet to be resolved is with IT security: end-to-end encryption may be required. If that is not possible, VoIP may reside on a separate voice network, which significantly degrades the business value but may improve voice quality.

Echo is a problem and difficult to diagnose: A considerable amount of time was spent diagnosing and fixing echo problems. Many possible sources/causes were identified, not all of which can be easily fixed. Trying to operate with mixed TDM and VoIP equipment makes it harder. Better industry-standard diagnostic tools would aid this process.

Headset selection is important: The appropriate headset selection is important for good voice quality. Wireless and less expensive wired headsets introduced noise and echo. The more expensive wired DSP headsets provided the best call quality.

G.711 Codec: G.711 is currently recommended for campus calls. Let gateways provide any necessary compression.

Drivers for VoIP include improved user mobility, functionality, presence, and productivity. Cost savings is on a case-by-case basis, with ROI and TCO improving as system costs decrease over time. In this case, the more specific business benefits we identified were as follows:

- Networks were simplified by converging voice and data networks.
- Costs related to data center space and telephone moves, adds, changes, and deletes were reduced.
- New or improved telephony applications were made possible due to VoIP's open, extensible architecture.
- The stage is set for advanced multimedia applications, services and productivity enhancements through converged communications.
- There is an immediate productivity gain from unified messaging, find/follow me, universal phone number, and presence integration.

These benefits can only be realized if acceptable voice quality is achieved.

ACKNOWLEDGMENTS

Thank you to the other members of the Intel VoIP Program Core Team for their work on this project and input to this paper: Dave Golish, Rick Lansford, Steve Sacker, Matt Santaiti, Rose Sheehy, Shelby Siegel, Geri Sunderlin, Gary Veum, Charlie Weaver, Mike Wilson, Michael Mall, Tracy Little, Roman Stobnicki, Scott Trevor, and Joe Walmach.

REFERENCES

[1] Session initiation protocol – RFC 2543, http://www.ietf.org/rfc/rfc2543.txt*

[2] Deploying VoIP – http://www.extremenetworks.com/libraries/solutions/DeployVoIP_WP.pdf*

AUTHORS' BIOGRAPHIES

Ranjan K. Sinha joined Intel as part of the Dialogic acquisition in 1999 as the Network/Telecom manager. He is currently responsible for the DMZ Operations at Intel.

He holds an M.S. degree in Computer Science from West Chester University, PA. His e-mail is ranjan.k.sinha at intel.com.

Catherine Spence is an IT computing manager for Intel. She has over 12 years experience with computer telephony and enterprise voice. She earned a graduate certificate in Software Engineering from Harvard University and a B.S. degree in Computers and Engineering from Trinity College, CT. Her e-mail is catherine.spence at intel.com.

Tim Verrall is a staff engineer with Information Technology. He is a lead architect and designer for Wide Area Network, network security, and voice networking at Intel. He has over 15 years of network engineering experience. His e-mail is timothy.verrall at intel.com.

Experiences with PC-Based Real-Time Multimedia Collaboration over IP

Shelby Siegel, Information Technology, Intel Corporation
Dave Lizotte, Information Technology, Intel Corporation
Blaine Bauer, Information Technology, Intel Corporation
Maria Frick, Information Technology, Intel Corporation
Duncan Glendinning, Mobility Group, Intel Corporation

Index words: VoIP, SIP, Wideband Codec, IP PBX, Peer-to-Peer, Conferencing, Multimedia

ABSTRACT

Intel recently completed a trial of standards-based real-time multimedia collaboration tools running on laptop computers over an Internet Protocol (IP) network. Key goals for the trial included validation of usage models and user benefits while utilizing the multimedia collaboration tools in a production work environment. Trial participants were equipped with a multimedia "softphone" application, a headset, and a Webcam that enabled them to establish high-quality small-group (multiparty) voice and video calls. Call setup was performed with the standard Session Initiation Protocol (SIP) and open-source products. These components provided a cost-effective, easy-to-setup and use collaboration environment, where all communications and collaboration were integrated into one device, the PC. This capability provided increased productivity because all information that people needed to do their jobs was literally at their finger tips.

The trial included an evaluation of high-quality voice calls using wideband codecs. These codecs encode twice the frequency range as that provided in the public telephone system, thus conveying most of the frequencies made by the human voice which enables them to deliver more lifelike speech and increased intelligibility. Because wideband algorithms are computationally intense, they most often appear in PC voice products. Thus the use of the PC for voice calls can deliver better quality conversations than we normally hear with telephones and cell phones.

The trial also tested PC client-based voice and video conferencing provided by the softphone application. Client bridging provides an easy way to conduct meetings on demand without requiring scheduling of conferencing bridges. The trial showed that laptops based on Intel® Centrino® mobile technology deliver the performance to conduct small group meetings of up to the limits of the "softphone" application (5 video or 7 audio-only participants).

In the trial, users found value in the integration of their communications environment into the PC, and in fact were looking for integration beyond the experience we provided. However, we also learned that both the tools and the infrastructure will need to be further optimized before users would be willing to fully replace their time-tested desk phone with a PC-based collaboration solution.

In this paper we describe the architecture, deployment, and key learnings for the trial. We also portray the business value of wideband audio and peer-to-peer conferencing.

INTRODUCTION

Collaboration capabilities are a key element of Intel's corporate growth strategy, as many employees participate in teams which span the globe. The ability to allow individuals to effectively collaborate remotely from their home sites as virtual teams, rather than physically attending meetings, allows employees to be more efficient with their use of time, and allows the corporation to reduce expenses by avoiding unnecessary travel. Many Intel employees are mobile–moving between conference rooms, cafeterias, labs, and their cubicles within their home campuses, working from home, traveling to other Intel campuses as well as visiting customer or supplier sites. Several years ago, Intel Information Technology (IT) launched a program to enhance employee productivity by equipping most Intel employees with notebook computers. Doing so allowed employees to carry with them the data and applications necessary to do their jobs. Over the past three years, IT has enhanced access to corporate data by equipping all Intel sites with wireless local area networks to enable access from anywhere within a site. IT continues to enhance the notebook computing environment by improving connectivity and collaboration tools to increase employee effectiveness.

Last year Intel launched a Digital Office initiative, which focuses on enabling real-time business by applying technology in four key focus areas for employee and IT manager benefit. The four key areas or "pillars" are 1) Seamless Collaboration, 2) Pervasive Connectivity, 3) Embedded IT, and 4) Information Assistance. Seamless Collaboration allows an employee to collaborate anywhere with anyone to accomplish his/her task. Since the Seamless Collaboration pillar highly correlates with the way Intel employees work, the Digital Office and IT organizations jointly defined a proof of concept trial to validate several key capabilities of Seamless Collaboration, including the following:

- Integrated collaboration consisting of voice telephony and client-to-client video, all running on the trial participant's notebook computer.
- Client-based (notebook computer) small group conferencing allows small group audio and video meetings conducted just on the laptop (with one laptop mixing the multiple audio and video streams), with no requirement to have other conferencing equipment nor to schedule the meeting ahead of time.
- "Click to call" capability uses a directory on the PC, which gives the user the ability to click on one or more names to instantly call people and establish a conference.
- High-fidelity audio conferencing, enabled with the use of wideband codecs on the PC.
- Unified messaging, the storage of voice mail messages in e-mail with the ability to listen to a message on the PC.

IT saw the trial as an opportunity to validate its strategic direction to transition multimedia calling to a Session Initiation Protocol (SIP)-based environment. Consequently, the trial was architected to include a SIP-compliant notebook softphone application (CounterPath Solution, Inc.'s eyeBeam* softphone [1]) along with the open source SIP proxy (SIP Express Router (SER) [2]) and PBX-gateway/unified message server (Asterisk* [3]) applications.

In order to focus the trial on the user experience and eliminate the potential of performance problems due to network congestion, the 100 trial participants were selected from four large U.S. Intel sites: Santa Clara and Folsom in California, Hillsboro in Oregon, and Chandler in Arizona. These sites are connected by high-bandwidth links with no congestion. Participants were asked to test the collaboration services while using wired connections (100 Mbps Ethernet) from their laptops to Intel's network. They also had the freedom to use the trial services while connected via 802.11b wireless LANs and via remote access over the Internet using VPN.

To prepare for the trial, a SER server was installed in Folsom and Asterisk servers were installed in the four U.S. sites to provide connections to the existing site TDM PBXs. Each participant was equipped with the eyeBeam client application, a headset, and a Webcam. A SIP dialing plan was created to give each participant the same number as their existing Intel phone number. This enabled trial participants to establish client-to-client IP calls in a way which is similar to normal Intel dialing, and to use the Asterisk gateways to make calls to people outside the trial. Furthermore, by forwarding their Intel phone to their local Asterisk server, every participant could receive calls from non-trial people on their eyeBeam client. Any incoming call that was not answered was handled by the Asterisk server, which took a voice mail message and routed it to the user's e-mail box.

The eyeBeam client includes several audio and video codecs. To enable the testing of wideband audio, the client was configured to use the open source Speex* codec [4] for calls between trial participants and the standard Public Switched Telephone Network (PSTN) G.711 codec for calls to people outside the trial.

Trial participants were selected as part of teams that frequently collaborate and were encouraged but not forced to use the capabilities exclusively. While it is understood that productivity benefits would increase with more consistent usage, the trial was specifically designed to evaluate how quickly and under what circumstances users would "take to" this new way of collaborating, rather than asking them to abandon their desk phone altogether.

Figure 1: The eyeBeam interface

They also were free to use the application in any way and for any purpose they chose (as long as the capability was enabled), while having access to a user guide with a detailed description of the features and capabilities. Feedback was collected via a Web page as well as a survey and selected interviews to correlate specific usage models with the user experience.

DETAILS OF TRIAL COMPONENTS AND DESIGN

Client

The eyeBeam PC softphone application (see Figure 1) provides a familiar telephone-like interface that is intuitive and can be used effectively with little training required. It provides many new telephony features such as “drag and drop” to place a call, auto answer, and auto conference. It includes basic telephony features, including hold and transfer, caller ID, speed dial, and microphone and speaker level controls. For the trial, a list of contacts (trial participants) was distributed to all participants in a format that could be imported into the client to facilitate speed dialing. The eyeBeam client also simplifies establishing small audio/video conferences (up to 7 users) with an easy-to-use interface. Users can be conferenced together at any time by clicking on the conference button.

VOICE AND VIDEO OVER IP

Voice over IP (VoIP) allows voice traffic to be integrated with an existing IP data network. This provides the advantage that a user can take their telephone with them anywhere as an application on their laptop. This also provides opportunities to integrate voice with other networked PC applications, such as presence, instant messaging (IM), and video.

VoIP uses codec technology to send speech across a network. When a person speaks, the sounds are picked up by the PC’s microphone and digitized and then encoded using a voice codec (coder/decoder), which is a piece of software integrated into a softphone or other VoIP device. At the receiving end another codec decodes the voice and converts it back to sound waves via a headphone or speaker.

VoIP audio and video calls can take place directly between distributed PCs (peer-to-peer) rather than through a central PBX. The voice codec can be negotiated on a per-call basis rather than being centrally determined; any codec that is supported on both ends can be used. This provides an opportunity to use a wideband codec, which encodes a much wider audio spectrum than a normal (narrowband) codec traditionally used in telephone systems (such as the PSTN). As shown in Figure 2, a wideband codec can capture and reproduce a frequency

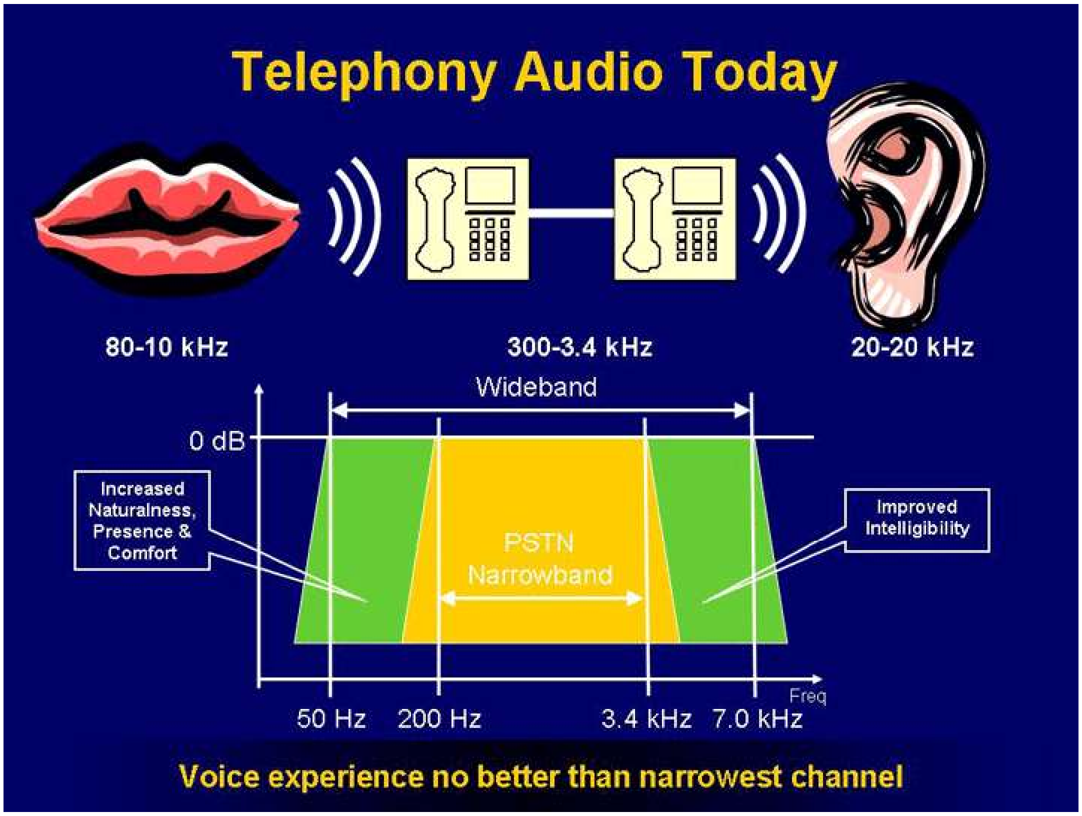

Figure 2: Frequency comparison

range of about 50 Hz to 7 kHz (out of the human vocal range of 80 Hz to 10 kHz) whereas a narrowband codec is limited to transmitting 200 Hz to 3.4 Hz, eliminating significant frequencies at both the bottom and top of the spectrum. Wideband audio brings telephone audio fidelity much closer to actual conversation–it sounds more like talking to someone face to face. The higher frequency range supported by wideband codecs can provide better understanding of intonations and accents and consonant discrimination (e.g., "s" versus "f"). This is becoming increasingly important as Intel's workforce and business becomes more global.

The eyeBeam client supports use of multiple codecs. High-quality wideband codecs can be used between eyeBeam clients while narrowband codecs are used to communicate with traditional telephones. In this trial, between eyeBeam clients, the Speex wideband codec is used; this provides very high-quality audio at a moderate 40 kbps bit rate (including IP overhead). Use of Speex also actually saves bandwidth compared to the standard G.711 uncompressed narrowband codec which is used in the public telephone network and which consumes at least 80 kbps (with IP overhead).

Video is provided using the standard H.263 codec at Common Intermediate Format (CIF) (352x288 pixels) and Quarter CIF (QCIF) (176x144 pixels) resolutions, by default implemented at up to 128 kbps. This generally provides a good picture in a window that is typically about 2 x 1.5 inches (depending on the monitor). The video codec is independent from the audio codec, so wideband audio can be used during audio-video calls.

One of the most significant factors in the quality of voice and video are the devices used to acquire the signal–a headset and camera. High-quality, moderate-cost devices were recommended for all trial participants to provide a baseline for signal acquisition.

Ensuring call quality with VoIP can be challenging, particularly on WAN links, and the technology is still being developed. For these reasons, the trial was focused on four US core sites, which are interconnected with unconstrained WAN bandwidth links.

Session Initiation Protocol

SIP is an evolving standard signaling protocol from the Internet Engineering Task Force (IETF) for Internet conferencing, telephony, presence, events notification, and IM. SIP is independent of the packet layer and only requires an unreliable datagram service, as it provides its own reliability mechanism. SIP is used for all signaling in this trial.

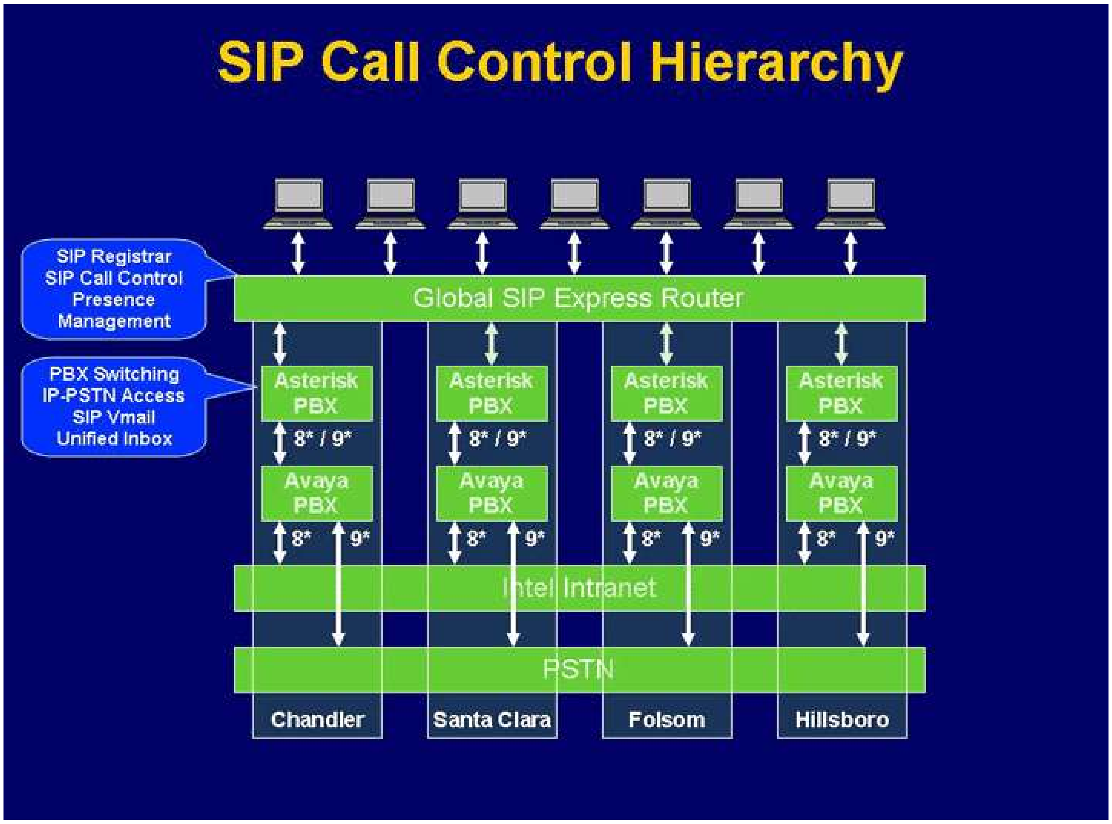

Figure 3: Trial architecture

The basic functionality for call setup is provided by a SIP registrar and proxy. The registrar/proxy for this trial is the open-source SER, which runs on Linux*. All users register with a single SER server, automatically providing each user a SIP address similar to an e-mail address. End users need only use the proxy as the SIP host; the proxy will use the PCs' IP addresses (provided through registration) to send the calls to the end PCs (eyeBeam clients). This allows a consistent SIP address regardless of a user's location; calls can be received on a single phone number anywhere in the world, regardless of one's physical location.

Since the only functions on the SIP proxy server are to register clients and proxy calls to end nodes, a single SIP proxy has minimal system requirements (less than 1 MB) and can scale to thousands of calls per minute. Multiple proxies could be used to establish SIP domains and redundancy, but for this simple trial only one proxy was required.

* Other names and brands may be claimed as the property of others.

Asterisk PBX

The trial also incorporates the open-source Asterisk PBX, running on Linux. Asterisk supports a long list of call features (see http://www.asterisk.org/features*), but in this trial is primarily used for voicemail, voicemail to e-mail integration, and as a SIP gateway to the TDM PBX infrastructure.

Unlike SER, which directs calls between IP end points (PCs), Asterisk can also terminate calls from the PCs (for voicemail and gateway functionality). The codecs that can be used are limited to those supported by Asterisk. These are typically narrowband codecs, which have lower CPU overhead. In the trial, G.711 is used between the client PC and Asterisk.

Overall Trial Design

The flexibility of SER and Asterisk provided numerous design possibilities. The trial design (see Figure 3) incorporated centralized call control and distributed voicemail and gateway functionality. This allowed all users to register on a single server (with a consistent client configuration), and still use a dialing plan similar to their existing TDM one.

Each of the four trial locations (Chandler, Folsom, Oregon, and Santa Clara) has an Asterisk server. In Folsom, the Asterisk server is running on the same system as SER. Each Asterisk server connects into a local PBX using a T1 ISDN line, each capable of 23 simultaneous calls. Two of the larger campuses involved have over 60 users. Traffic studies indicate that the Asterisk gateways to the PSTN were utilized 25% on average with peak traffic hitting 50-70% of trunk utilization. During some peak gateway testing a noticeable decrease in call quality was observed. The cause appeared to be isolated to the gateway card technology and the amount of processes being utilized on the server to support the gateway card. One possibility to resolve this call-quality degradation is to ensure that each card function runs on a single process.

Dialing to regular phones is provided by the local Asterisk server at the trial locations. To keep the dial plan consistent for end users, the trial used the same dial plan as Intel's existing phone system.

ENGINEERING KEY LEARNINGS

One of the main goals of this trial was to increase the knowledge of the IT network engineers on the topics of SIP, open source, and VoIP support issues. The trial was very successful from this perspective and produced key learnings in the areas of open-source software, SIP interoperability, VoIP monitoring and measurement, and client configuration issues.

Use of open-source software has both advantages and challenges. The upfront cost and source code access is seen by many as an advantage. The lack of a formal support model is seen by some as a disadvantage, and in this trial resulted in some implementation delays. Rather than direct support, most problems must be solved by sifting through discussion archives on the Internet. Some of the developer solutions that were utilized in the design did not function as the documentation indicated, and further research and experimentation were needed to resolve these feature/functionality issues.

This experience produced some guidance for future projects. Extra time should be allocated for lab testing before the proof of concept and design changes should be anticipated. Engineers can learn a considerable amount through trial-and-error lab testing, but if an organization wishes to initiate a proof of concept quickly or it does not have network and telecom engineers with the technical experience to resolve issues independently, it should consider employing an outside consultant with open-source experience.

There appears to be a significant opportunity for companies to pre-package/bundle open-source VoIP solutions in much the same way Linux distributions are now bundling solution stacks. Open source VoIP technology is a disruptive yet driving force in the VoIP PBX industry. Open source leverages the collective efforts of many developer groups and technical experts to create new technologies and applications. With the high demand for telephony customization and application integration, open source will continue to drive vendors to either embrace it or stay ahead of its development.

The trial showed that basic SIP functionality was interoperable among a variety of products. However, the basic IM and presence functionality proved more problematic. Support for the IM and presence capability of the softphone client was part of the initial plan, but presence management was not reliable with the version of SER used in the deployment, so the trial was rolled out without this capability.

Monitoring and measurement, particularly call detail recording, is a challenge with open source IP PBX solutions. The data files were collected during the trial, but the information provided was very limited. Further tool integration or software is required to manipulate the data records into a meaningful reporting format. Though there are numerous open source tools available for SER and Asterisk, none were integrated into this trial environment.

Client stability is important when calls are made through the PC, and this provided some challenges due to the interactions of the real-time communications capabilities with various IT-mandated client applications (such as virus detection, firewall, and remote management). This trial was designed to evaluate the usage of a real-time application in an unconditioned environment, so the clients and network were not tuned to support this usage model. One of the key learnings was that the introduction of telephony software as a real-time collaboration tool on the PC requires characterization of the entire PC software environment to ensure coexistence of all applications before global deployment of a fully supported solution can be attempted.

USER EXPERIENCE

This trial was designed primarily to establish the value of "seamless collaboration," as opposed to a product-centric proof of concept where the focus is on the tools themselves. To this end, the users were requested to use the eyeBeam client as their primary communication device for all their audio collaboration needs, but not given any specific scenarios to test. Wherever possible, intact teams that normally collaborated with each other were selected for the trial, so they would get the most out of the solution, and so there would be a high likelihood of peer-to-peer video and telephony calls utilizing wide band audio.

However, from a user perspective, it becomes difficult to separate the value proposition from the day-to-day issues encountered with the way the tools work. The user feedback falls mainly into the following categories:

1. Understanding the capabilities and functionality of the solution.
2. Desired functionality and usability issues with the interface.
3. Feedback on technical limitations and audio quality.

The users were provided with a detailed Intel-specific user guide explaining the setup and features of the eyeBeam client at the start of the trial, along with a link to a FAQ page, a support email account, and a Web page where they could post their feedback. The FAQ page was updated regularly throughout the trial based on the feedback received.

Most of the issues encountered during the first couple of weeks were related to the setup of the eyeBeam client, audio tuning, and volume control (such as use of the headset, mixing of PC and phone audio, and coping with different levels of volume for different participants on the call). Most users would like to see a client with auto configuration so that manual set up is not required to start using the technology. It will be very important to support an automated installation and configuration process for large global deployments. Further development work with vendors' developer tool kits can assist to minimize user setup time, configuration errors, and trouble calls upon initial setup.

Redirecting all incoming calls to the eyeBeam client rather than the desk phone proved a bit challenging as well. Users were given the opportunity to utilize their existing desktop phone or the proof of concept technology on their laptop. This architecture allowed for minimal productivity loss if issues arose with one or the other technologies.

Users commented a lot on the dialing functionality, ranging from requests for e-mail integration ("click-to-dial" a name) to drag-and-drop, cut-and-paste and number editing functions, to voice recognition similar to cell phone functionality. Some would have liked a quick access toolbar in other office applications, rather than having to switch to the eyeBeam application to initiate the call. Several users would have liked either IM functionality within eyeBeam or at least strong integration with Intel's existing IM application.

Answering an incoming call via eyeBeam also took some getting used to. While the application always pops to the front when a call is received, in some cases users experienced "stolen keystrokes" which produced unpredictable results within eyeBeam or promptly sent it to the background again while typing in other applications. This lead to some missed calls. Since all audio goes to the headset, users also felt they were at risk of missing calls if they did not wear the headset all the time at work.

Many users compared the eyeBeam client to their experience with other softphones. In some cases, the comparison was in favor of eyeBeam, in others less so. Some found the user interface too complex and were not able to navigate the different buttons, others got hooked on the AutoAnswer and AutoConference features and actually changed the way they scheduled meetings as a result.

At the same time, users commented that the ability to make a call by a simple drag and drop of a name, the ability to click-to-dial, and the autoconferencing features are a great improvement over desk phones. These softphone features provided them increased productivity.

We received a wide range of feedback on audio quality. Several users heard significant pitch changes at times during calls when there seemed to be resource contention, either because they were using other applications at the same time, or because they were making calls over a VPN connection, from a wireless hotspot or their home office. For some users calls over their VPN broadband connections were of good quality while the experience for other users over VPN broadband rendered the client useless. During a call, when the PC was performing a resource-intensive task, such as when a large e-mail attachment came in over VPN or when opening a large presentation, a slowdown of the desktop and changes in voice quality were evident. These were considered annoying but not detrimental to productivity.

At the same time, all users that experienced true wideband (eyeBeam to eyeBeam) calls reported amazement at how much clearer and easier to understand the audio was compared to their regular desk phones. For some, it even sounded better than other Internet telephony tools. Given that many users still dialed into a standard audio bridge for their conference calls, this is one of the benefits that can only be fully realized if a large number of users has access to and is utilizing the technology.

Conversely, the Voice Activity Detection (VAD) feature, which sends "digital silence" when there is no voice detected, was unexpected and confused users. To many trial participants it seemed to be a loss of audio or a dropped call because it was too quiet. Once users turned off VAD there tended to be more "static" on the line but the experience was more familiar. Eventually, softphone developers will incorporate side tone, which will ameliorate this issue.

Advanced technical users suggested that eyeBeam should provide diagnostics on packet error rate, jitter, and latency so the user can "monitor" the health of the conversation and adjust as needed. Being told that the other party could not hear them certainly was perceived as a negative, even though in most instances this was related to a weak signal from a Bluetooth* headset rather than the inherent audio quality in eyeBeam itself. Those users also would have liked to have a call automatically reconnect itself once the signal from the Bluetooth headset was within acceptable range again.

Headsets in general drew some interesting comments. Most users liked the quality they got via the recommended headsets but found them too bulky to travel with. Headsets utilizing the audio jack on the PC did not produce the same high fidelity. The inbuilt speaker and microphone on the PC were not attractive due to the lack of sufficient echo-cancellation software and array microphones. Not being able to accept a call directly on a Bluetooth headset also was a usability issue.

Little feedback was received about the video. Call detail records on the server showed that very few video calls were actually established, indicating that many users were either not comfortable using video, could not be bothered to try it, or did not have the time or patience to figure out how it works. Some users did feel the video aided comprehension in remote one-on-one meetings.

Unexpected was the fact that sometimes calls seemed to go into "dead air"; this was traced to a user initiating PC hibernation without logging off of eyeBeam. Closing the eyeBeam client and therefore unregistering the phone prior to going into hibernation mode was an easy enough work-around but needed to be explained first.

Regrettably, many users also experienced hard lock-ups on their PCs while using the softphone client, which discouraged usage to a certain extent. The hard lock-ups were traced to an interaction with an IT-mandated application load, which similarly affects other real-time-softphone and collaboration clients. We recommend characterization of the current PC operating system (OS) and any additional software components before deploying real-time applications into production.

In general, traveling users seemed to get most value and benefit out of the tool since their desk phone essentially followed them. Numerous user anecdotes were received regarding the value of the softphone in environments where there was no corporate phone system phone (such as a lab, hotel room, or phoneless cubicle). Also, it made settling down in a remote campus very easy, especially given the lack of office space and traveler workspaces equipped with phones at some Intel sites. Being able to receive their voice mail directly in e-mail contributed to this positive experience and was perceived as a significant value-added feature.

Generally, the comments were encouraging and confirmed that employees can be more productive with real-time collaboration tools on the PC since there are incremental time savings in many repetitive tasks, as well as increased support for mobility and improved comprehension in conversations.

CONCLUSION

The user trial was conducted for a period of about two months. It revealed that users found value in the integration of their real-time communications with their PC collaboration environment, but wanted even more integration, such as with their e-mail client and IM. Users particularly enjoyed the wideband audio experience, the "click to call" capability, unified messaging, and the ability to be mobile and still communicate from a variety of locations. However, they also were bothered by the inconsistency of the experience, as the trial included no mechanisms to ensure that the real-time communication was given priority on the PC and the network. Video was not used very much, even though some value was noted for one-on-one meetings. Finally, the trial showed that some SIP products interoperate, but it takes work (especially with open source) and possibly multiple products to obtain a desired feature set.

Overall, the user feedback was positive on the added mobility and productivity enhancements this PoC provided. IT's plans are to further develop the integrated communications and collaboration environment with the PC, taking advantage of the advanced collaboration features that will be delivered with the next-generation laptops based on Intel Centrino mobile technology.

ACKNOWLEDGMENTS

The authors recognize the contributions of other members of the project team, including Tim Verrall, Darrin Donithorne, Dave Wescott, Greg Trusley, and Gary Veum. We also thank MPG for their support of this project.

REFERENCES

[1] eyeBeam soft phone from CounterPath Solutions, Inc., http://www.xten.com/index.php.*

[2] SIP Express Router (SER) from iptel.org, http://www.iptel.org/ser/.*

[3] Asterisk from Digium, http://www.asterisk.org/.*

* Bluetooth is a trademark owned by its proprietor and used by Intel Corporation under license.

[4] Speex from Xiph.Org Foundation, http://www.speex.org/.*

AUTHORS' BIOGRAPHIES

Shelby Siegel is an architect for VoIP and real-time communications in Intel Information Technology. He joined Intel in 1975, where he has worked in information technology on operating systems and networking. He received a B.S. degree in Mathematical Sciences and an M.S. degree in Computer Science: Computer Engineering from Stanford. His e-mail is shelby.siegel at intel.com.

Dave Lizotte is a senior voice engineer at Intel Corporation in Folsom, California. He is currently focused on new capabilities development for Voice over IP and convergence solutions. Dave studied Telecommunications and Networking in Washington State, and received a Telecommunications Engineering degree at J.M. Perry Technical Institute. Dave joined Intel in 1994 and can be reached at dave.a.lizotte at intel.com.

Blaine Bauer is a senior network engineer focused on developing new capabilities for quality of service, rich media distribution, network management, and converged communications. Blaine has a B.S. degree in Computer Science from Kansas State University and a Masters in Systems Management from the University of Southern California. He has been involved in TCP/IP networking since 1986, is a 10-year Cisco Certified Internetworking Expert, and has been at Intel since 1995. Blaine can be reached at blaine.d.bauer at intel.com.

Maria Frick is a project manager focused on Real-Time Collaboration projects in Intel Information Technology. She has a Bachelor's Degree in Languages from the Institute of Foreign Languages in Munich, Germany, and has spent most of her career in the field of Localization and Internationalization. She joined Intel in 1997 and can be reached at maria.frick at intel.com.

Duncan Glendinning is an architect in the Mobility Group Strategic Planning organization, where he is focused on communication-related topics for notebook and handheld computers. He received his B.Eng. and M.Eng. degrees in Electronics both from Carleton University. Duncan joined Intel in 1995; he can be reached at duncan.glendinning at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

Using Intel® Technologies to Build Next-Generation Media Servers

Joseph Grecco, Digital Enterprise Group, Intel Corporation
Mark Mize, Digital Enterprise Group, Intel Corporation
Ranjan Singh, Digital Enterprise Group, Intel Corporation

Index words: Media Servers, Video, VoIP, IMS, NGN

ABSTRACT

In this paper we describe the role of the media server in traditional Public Switch Telephone Network (PSTN), Internet Protocol (IP), and next-generation networks. We examine the Intel® building-block technologies and their use in developing powerful, cost-effective multimedia communication solutions. We begin with a description of the components that make up a media server and then outline three representative network configurations into which media servers are deployed. The following sections examine the Intel technologies used to build a next-generation media server from the top down. We start by taking a look at Intel NetStructure® Host Media Processing (HMP) Software, and how it makes developing interactive multimedia applications straightforward and cost effective; we walk through the steps required to build a simple voice and video application. The foundation of Intel NetStructure Host Media Processing Software is the Intel Architecture processors, chipsets, and Intel® Integrated Performance Primitives (Intel® IPP). We show how Intel Architecture processors and Intel IPP can offer world-class performance of media processing algorithms such as audio and video codecs. Finally, we look at how developer tools like the Intel C++ Compiler and VTune™ Performance Analyzer can be used to produce high-performance application code.

INTRODUCTION

Fixed-mobile convergence–the convergence of wire line and wireless devices into a single telecom network–promises to enable service providers to reach a greater number of potential customers with a wider range of service offerings. Service providers in the telecommunications industry are looking to multimedia to increase revenues by allowing them to add video to their traditional voice services, thus enriching the end-user experience and making the services more attractive to their subscribers. For example, adding video content to a traditional voicemail application enables a new level of personalization. Users can be greeted with a subscriber's personal video, humorous clip, or cartoon animation. This level of personalization is particularly appealing to the teenage subscriber base.

Greater levels of personalization are being provided via new revenue-generating services such as video color ringback and video caller ID. For example, with video color ringback the subscriber can replace the standard ring tone that the caller hears with a personalized video message. Imagine seeing and hearing the person you are calling running toward the phone yelling, "I'm coming…." IP Multimedia Subsystem (IMS) from the Third Generation Partnership Project (3GPP) defines media servers as part of its next-generation network architecture for multimedia solutions. Similarly, in the enterprise, multimedia solutions are being investigated in conjunction with IP technology as a means of reducing operating expenses while improving worker productivity and customer satisfaction.

Intel NetStructure Host Media Processing Software simplifies the development of multimedia telecommunication applications: supporting new capabilities such as video is simple and straightforward.

The move toward IP-centric solutions can represent significant changes in the way media servers and applications are modeled. Intel is committed to protecting our customers' investments. Existing applications that run on Application Programming Interfaces (APIs) delivered with Intel NetStructure Host Media Processing Software can easily migrate from a traditional PSTN environment to an IP-based next-generation network. At the same time, Intel embraces emerging and evolving open industry standards in architectures, protocols, and interfaces.

The powerful combination of Intel Architecture processors and chipsets, innovative platform technologies such as dual core, Hyper-Threading Technology[1] (HT Technology) and Intel® EM64T[2]; and software technologies such as Intel IPP and Intel NetStructure Host Media Processing Software, will meet the demanding performance requirements of media signal-processing applications. By supporting these technologies with a wealth of available development, test, and performance-tuning tools, Intel offers an exceptional signal-processing platform for building next-generation media servers with the lowest total cost of ownership.

As new and innovative technologies such as multi-core and Intel Virtualization Technology become available, Intel platforms will provide more end-user value through new features and improved performance.

These Intel products and technologies are used to deploy services in a media service network. A media service network is a network through which media services are provided to an end user. We start our discussion with a description of the components that make up a media service network using three representative network configurations. We then take a top-down approach to discussing the Intel technologies used to build a next-generation media server. We start by taking a look at Intel NetStructure Host Media Processing Software, and how it makes developing interactive multimedia applications straightforward and cost effective. We walk through the steps involved in developing a simple voice application using Intel NetStructure Host Media Processing Software APIs. We then examine what is involved in adding video capabilities to the application, demonstrating how developers can add features to existing applications while retaining their investment in existing code.

At the heart of an Intel-based media server is the Intel Architecture processors, chipsets, and Intel IPP. In the subsequent sections we show how these technologies can offer outstanding performance of media processing algorithms such as codecs. In the last sections we look at how developer tools like the Intel C++ Compiler and VTune Performance Analyzer can be used to produce high-performance application code.

TAXONOMY OF A MEDIA SERVICE NETWORK

A media service network is a network through which media services are provided to an end user. The services are implemented by applications controlling media resource functions on a media server. In this section we discuss the components that make up a media service network, illustrate three representative communication network configurations into which media servers are deployed, and describe the interfaces used by media service applications.

There are a number of architectures in use today that define the composition of a media service network. Some are specified by industry standards bodies and others are proprietary. Most can be generalized to a simple model where the media server interoperates with a number of other components to form the complete solution.

Components of a Media Service Network

We will first define the basic components of a media service network: end-user terminals, application server, media server, media store, and gateways.

End users access media services through the public network from a local device we will call a terminal. Examples of terminals are a plain old telephone (POT), a video-enabled cell phone, and a multimedia PC.

The application logic that realizes the media service is hosted by an application server. The application server provides the runtime execution environment for the application.

Media processing such as transcoding and video image processing (e.g., text overlay, resizing) are performed by media processing resources on the media server.

The media store is responsible for the storage and retrieval of the media (i.e., to/from disk).

---

[1] Hyper-Threading Technology requires a computer system with an Intel® Pentium® 4 processor supporting HT Technology and a HT Technology enabled chipset, BIOS and operating system. Performance will vary depending on the specific hardware and software you use. See www.intel.com/homepage/land/hyperthreading_more.htm for additional information.

[2] Intel® EM64T requires a computer system with a processor, chipset, BIOS, operating system, device drivers and applications enabled for Intel EM64T. Processor will not operate (including 32-bit operation) without an Intel EM64T-enabled BIOS. Performance will vary depending on your hardware and software configurations. See http://developer.intel.com/technology/64bitextensions/ for more information including details on which processors support Intel EM64T or consult with your system vendor for more information.

protocols are in use), the media server will have to interoperate with gateways that are responsible for translating between the different network protocols. Gateways are typically classified as signaling gateways that translate one signaling protocol to another and media gateways that translate from one media format to another. We explain signaling in the following section.

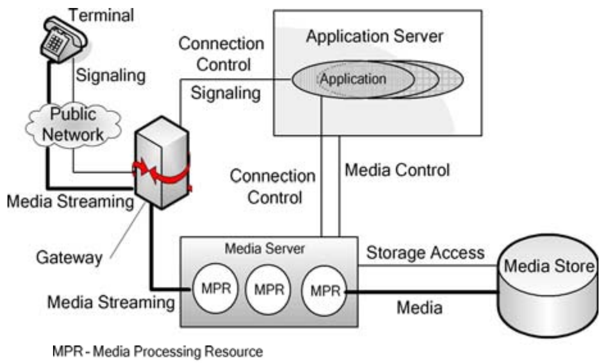

Figure 1: Basic media service network architecture

Control and Data Plane

Components in a media service network can be thought of as interoperating in two planes. The data plane, where the actual media traffic flows, is depicted as the thick lines in Figure 1. The control plane, shown as the thin lines, is where control and signaling messages flow. Signaling messages are used to establish and manage the media session between the end-user terminal and the media server. The session contains the state and any other context associated with the connection. The terms "session" and "call," as in a telephone call, are used interchangeably. Signaling typically originates from the end-user terminal and the application, potentially being translated by a signaling gateway. As mentioned, the media traffic flows along the data plane. Routing and flow of the media are controlled using a connection control protocol in the control plane. Connection control typically originates from the application and is directed toward the terminal or gateway and media server.

The application logic controls the media server through a media control API, discussed in more detail later.

Composing a Media Server

The media server is the architectural element responsible for transmitting media content to an end user over a communications network. The content may be in the form of video, audio, text, or a combination of the three. The media server often contains digital signal processing resources to process or transform the media prior to transmitting it to the user (e.g., gain and speed control of an audio stream). In addition to transmitting media content to the end user, a media server may also be capable of accepting control input from the end user's terminal. This input may be used by the media service application, making it interactive.

Figure 1 illustrates the logical components of a media service network. In this diagram the application server, media server, media store, and gateway are shown as separate entities. This is a logical organization and does not necessarily imply a physical decomposition. In practice, depending on the individual system requirements, any or all of these elements can be combined within a single node. In the following sections we will often simplify the diagrams by combining the application server, media server, and media store elements into a single node.

CIRCUIT-SWITCHED NETWORK

The traditional circuit-switched telephone network, also known as the Public Switch Telephone Network (PSTN), is a connection-oriented communications system where dedicated "circuits" are used to transport the media between the end user and the media processing resources. The PSTN consists of a collection of switches, as well as connections between switches and end-user terminals (a.k.a. telephones).

The establishment and routing of circuits between the end-user devices and the central office switches are controlled via signaling. Traditional phone devices use in-band signaling, where the signaling information is carried over the same circuit as the media. An example familiar to everyone is the generation of DTMF tones from the touch-tone pad of a common telephone used to identify an end user by number.

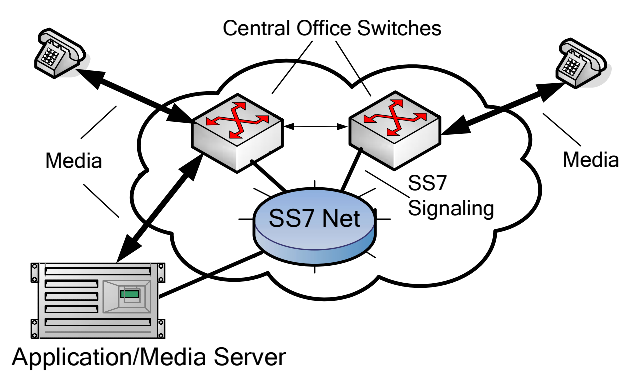

Figure 2: Traditional circuit-switched network

Signaling System 7 (SS7) [1] is an international standard signaling system that provides an efficient and protected method for routing circuits through the network. SS7 is an out-of-band signaling method that reserves dedicated circuits called signaling links for transmitting the signaling information. The Integrated Services Digital Network (ISDN) [2] also defines an out-of-band signaling protocol. In either case the control, or signaling, information is separate from the media.

Media servers can use in-band signaling to control the connections between end-user devices and media resources. However, out-of-band methods such as SS7 have become essential for many media service applications and have become the method of choice for media servers deployed today (see Figure 2).

Intel has a number of technologies that allow media servers to connect to the PSTN circuit-switch network using in-band or out-of-band signaling techniques.

PACKET-SWITCHED NETWORK

Unlike circuit switching, which relies on dedicated point-to-point connections to transmit contiguous streams of media, a packet-switched network breaks up the media stream into small message packets. Each packet contains address information specifying the desired destination for the packet. Because packets are addressed, no pre-established communication path or reserved "circuit" is required for a packet network. In contrast to a circuit-switched network, bandwidth in a packet network is used as needed to send packets. A quiescent channel need not produce load on the network.

As packets traverse the packet-switched network, they do not necessarily follow the same path to the destination. Network traffic conditions or outages can result in packets being dynamically routed through different paths. As a result, packets may take different amounts of time to reach the destination, they may arrive in a different order from the one in which they were sent, or they may be lost altogether. It is the responsibility of the destination device to deal with packet order, latency, jitter, and loss. Once they are reassembled, the destination device can render the media in real time. Packet loss, network latencies, and packet ordering represent some of the biggest challenges to providing reliable, high-quality, real-time communications over a packet network.

Several protocols have been designed for packet-based networks, the most popular being the ubiquitous Internet Protocol (IP). IP, which specifies the addressing scheme and packet format, was originally designed for data communications between dissimilar computers. It is a connectionless protocol and not inherently reliable. Numerous additional protocols have been designed to run on top of IP specifically to address the needs of real-time voice and video communications.

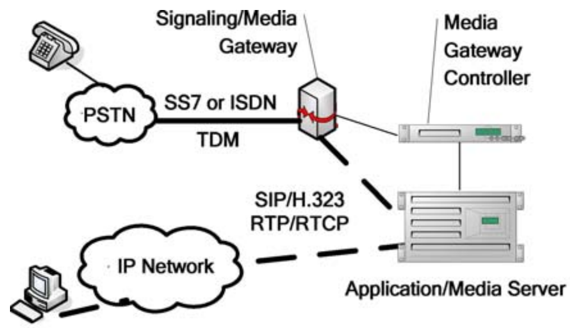

Figure 3: IP network with PSTN gateway

Two predominant standards have emerged in the IP signaling domain: H.323 [3] and SIP [4]. H.323 is an International Telecommunications Union (ITU) standard for the transmission of voice, video, and data over an IP network. An umbrella specification that consists of a suite of protocols and standards, H.323 defines a complete framework for multimedia communications. It specifies detailed protocols, messages, and state machines. H.323 has its roots in the traditional telephone network and attempts to address a wide range of problems including basic call states, supplementary services such as call forwarding and call waiting, Quality of Service (QoS), mobility (users moving from one address to another), and security.

SIP is a protocol defined by the Internet Engineering Task Force (IETF) that has begun to supplant H.323 in popularity. Unlike H.323, which attempts to address specific functionality such as basic and supplementary multimedia services, SIP defines a protocol that supports a generic session model upon which systems can be built. The SIP specification (IETF RFC 3261) defines a small set of messages that addresses location services, session creation and termination, and session parameter passing. It is designed to support a wide range of multimedia applications.

Although SIP and H.323 seem to have approached the problem of multimedia communications from different angles, they both define a number of architectural elements that enable location services, authentication, mobility, and interoperability with the circuit-switched PSTN. At a high level, SIP and H.323 are similar in terms of functional decomposition. In addition, both SIP and H.323 use RTP/RTCP [5] (IETF RFC 3550) as the media streaming protocol over IP.

For interoperability with the PSTN, SIP and H.323 define an architectural element called a gateway. The gateway is divided into two functional components. The signaling gateway (SG) component converts between PSTN control signals and the appropriate SIP/H.323 messages. The media gateway component (MG) converts between circuit and packet media. A MG controller is the entity that controls the MG (see Figure 3).

In architecture diagrams, the MG, SG, MG controller, application server, and media server are often shown as individual devices. Depending on the system requirements, these devices may be implemented on separate nodes or be physically combined within a single server.

3GPP and IMS

The Third Generation Partnership Project (3GPP) developed the IP Multimedia Subsystem (IMS) [6]. IMS is an example of a next-generation communications network architecture. It enables service providers to deploy new IP-based, multimedia communication services over both the fixed wireline and mobile telecommunications networks.

With IMS, services can be provided over any IP network (GPRS, WLAN, etc.). The IMS infrastructure is IP based, using standard SIP/IP between the core network elements. Originally designed for the mobile network, IMS can provide IP-based services to external circuit-switched networks as well as external IP networks.

The IMS architecture defines functional entities falling into the six main categories listed in Table 1. The IMS entities collectively address interoperation with other networks (e.g., circuit-switched and radio access networks), security, roaming, policy control, billing, and service deployment.

Figure 4 shows a simplified view of the IMS architectural elements that cover session management and call routing, security and policy management, network interoperability, security, and services.

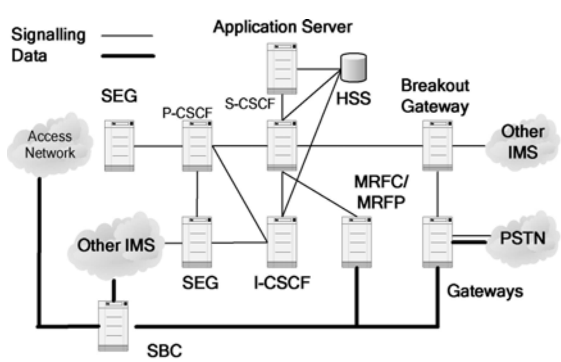

Figure 4: Simplified IMS architecture

Table 1: IMS functional categories

Session and routing	Call Session Control Functions (**CSCF**)
Databases	Home Subscriber Server (**HSS**), Subscription Locator Function (**SLF**)
Interoperation	Breakout Gateway Control Function (**BGCF**), Media Gateway Control Function (**MGCF**), IMS Media Gateway (**IMS-MGW**), Signaling Gateway Function (**SGF**)
Services	Application Server (**AS**), Multimedia Resource Function Control (**MRFC**), Multimedia Resource Function Processor (**MRFP**)
Support	Policy Decision Function (**PDF**), Security Gateway (**SEG**), Topology Hiding Inter-network Gateway (**THIG**)
Charging	Charging Collection Function (**CCF**)

The IMS elements most relevant to this paper are the Multimedia Resource Function Controller (MRFC), Multimedia Resource Function Processor (MRFP), and the Applications Server (AS). The MRFC is the element responsible for taking SIP requests from the AS and translating them to messages that control the media processing resources residing in the MRFP. The MRFP is where the actual media processing resources reside.

While the IMS specifications separate the AS, the MRFC, and the MRFP, implementations can combine one or more of these elements into a single node, as noted earlier. In fact, it is widely expected that the MRFC and MRFP elements will typically be deployed as a single unit.

In Figure 5 we show a combined AS, MRFC, and MRFP and collapse the other IMS elements. This diagram shows many of the same functional elements as shown in Figure 3: an IP-based Media Server interoperating with the circuit-switched network via a media and signaling gateway combination. In fact, many of the same Intel components may be used to build both systems.

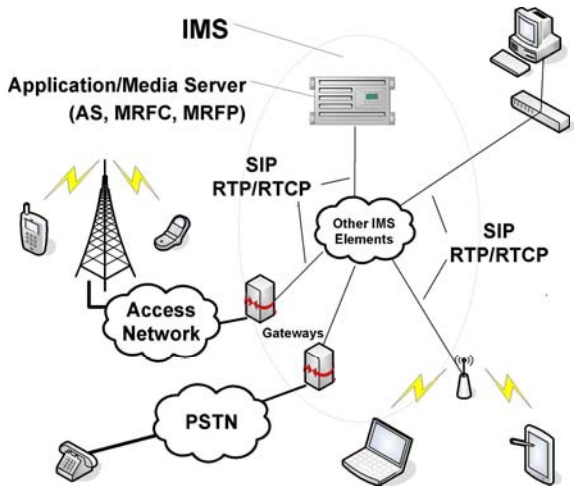

Figure 5: Media server view of the IMS architecture

APPLICATION PROGRAMMING INTERFACES

The components shown in Figure 1 interoperate along the control plane through a collection of interfaces. Of particular interest to the media service application developer are the interfaces between the application server and the media server. These APIs can be divided into three categories: signaling, connection control, and media control.

As was previously mentioned, signaling messages are used to establish and manage the media session between the end-user terminal and the media server. The primary signaling interfaces for next-generation applications are SIP for IP and IMS and Signaling System 7 (SS7). Intel NetStructure SS7 Boards provide applications with basic and advanced signaling capabilities in both wireline and wireless networks. For SIP signaling, applications can use either SIP protocol stacks from third parties or Intel's own APIs. Intel NetStructure Host Media Processing Software supports standard SIP protocols through the Global Call API. This is discussed in more detail below.

Connection control APIs allow the application to control the routing and flow of media along the data plane. There are numerous methods for controlling the media data flow. For the IP-centric next-generation media server, the most common method involves directing Real-Time Protocol (RTP) streams using SIP. The Media Gateway Control Protocol (MGCP) [7], as defined by the Internet Engineering Task Force (IETF), allows applications to control the routing and flow of media through a gateway.

The category of interface most relevant to this discussion, media control, covers a wide variety of media functions. It includes basic voice messaging such as the ability to play announcements and record messages, as well as video functions. Media control also covers the ability to accept control input from the end user. This input may be in the form of in-band tones such as DTMF or speech, as well as out-of-band messages such as those defined by RFC-2833 [8]. Intel supports a variety of APIs for media control. In the following paragraphs we discuss a few notable ones.

Voice Extensible Markup Language (VoiceXML) [9] is a standard XML-based language that enables voice/speech interactions with Web pages. VoiceXML is defined by the World Wide Web Consortium (W3C). Application logic is expressed in the form of VoiceXML documents that are processed by a VoiceXML interpreter (a.k.a. browser) which, in turn, drives the media server. VoiceXML is popular for Interactive Voice Response (IVR) applications because it defines a simple yet powerful high-level programming language.

Two additional media control interfaces attracting interest for next-generation media servers are the Media Sessions Markup Language (MSML) [10] and Media Objects Markup Language (MOML) [11]. MSML and MOML are Internet drafts submitted to the SIPPING work group of the IETF. MSML/MOML are XML-based interfaces that, while capable of standing on their own, are designed to complement one another. In contrast to VoiceXML, which presents a high-level language, MSML/MOML defines a framework for describing a wide range of media functions in an extensible manner. Besides offering rich semantics, MSML/MOML is attractive for use by next-generation media servers because it is designed to be compatible with SIP. The downside of MSML and MOML is that they are not yet recognized as standards and are still relatively immature.

Another notable media control interface is Intel's R4 "C" language APIs. The R4 APIs have been a staple of computer telephony applications programming for over 10 years. They are supported on Intel media processing technology boards as well as Intel NetStructure Host Media Processing Software.

All of the interfaces described above can be used with Intel-based media servers. Some, such as VoiceXML, can be acquired through third parties, while others are directly available from Intel. In the following sections we discuss Intel NetStructure Host Media Processing Software and the R4 APIs. We demonstrate how to implement a simple media service application that will run on media servers built using HMP software.

INTEL NETSTRUCTURE® HOST MEDIA PROCESSING SOFTWARE

Intel NetStructure Host Media Processing Software is an Intel architecture-based technology for processing media.

It is an ideal platform for implementing a media server, whether in the rigorous environment of an IMS network or in less standardized applications.

About Intel NetStructure® Host Media Processing

Intel NetStructure Host Media Processing Software runs on a general-purpose computer running Windows* or Linux*. The HMP software processes media on the Intel CPU and interfaces with external entities using standard IP protocols through the computer's Network Interface Controller (NIC).

Intel NetStructure Host Media Processing Software supports all common media server functions, including T.38 fax, audio/video play, audio/video record, audio conferencing, audio transcoding, IP call control, and also provides hooks for speech recognition and text-to-speech. As of this writing, HMP software from Intel runs up to 400 concurrent channels (i.e., bi-directional audio sessions) on a single computer. Additionally the HMP software interfaces with PSTNs, PBXs, and digital stations through telephony interface cards [13,28]. Table 2 lists the features in more detail.

Table 2: Intel NetStructure Host Media Processing Software feature tables

Network Interface	
	IP over Ethernet, Traditional Telephony interface cards

IP Call Control	
Protocol	H.323, SIP
Integration with third-party call and connection control stacks	Provided via R4 IPML

IP Streaming	
Protocol	RTP (RFC 3550)
Audio formats	G.711 A-law, µ-law 8-bit 8K G.723.1 (5.3 KHz and 6.3 KHz) G.729a, G.729b
Video formats	H.263 profile 0, level 30 (RFC 2190)
QoS	Alarms Frames per packet control Packet loss reduction RTP/RTCP timeouts
Tone generation and detection	RFC 2833 H.245 User Input Indication
Media control over RTP	Programmatic control of inbound RTP stream gain and output RTP stream volume

Voice Processing Features	
Features supported	Play, record, and tone generation and detection
Play	Volume control and index play
Record	AGC
Audio file formats for play/record	OKI ADPCM 24 Kbps, 32 Kbps G.711 A-law, µ-law @ 48 Kbps, 64 Kbps All of the above in WAV format Linear PCM 8b @ 11 Khz (Wave format only) Linear PCM 64 Kbps Linear PCM 128 Kbps G.726

Conferencing Features	
Total parties/server	240
Advanced Features	N-way summing Coach/pupil mode DTMF detection DTMF clamping Active talker notification

User Input Features	
In-Band	In-Band DTMF generation and detection User-defined global tone generation and detection (GTG, GTD)
Out-of-Band	RFC 2833 RTP Messages H.245 User Input Indicator

* Other names and brands may be claimed as the property of others.

Video Processing Features	
Play	Playback of synchronized voice and video
Record	Stores synchronized voice and video to file
Video format	H.263 (profile 0 level 30)
Picture Sizes	CIF, QCIF, and sub-QCIF
Video file formats	Dual proprietary file formats Audio file (.pcm): Linear PCM 16b 8K Video file (.vid): H.263 bit-stream data
Offline conversion tool	Convert AVI Type-2 (DVSD or DV25) files (PAL or NTSC) to proprietary format Convert proprietary format to 3GP Release-4 file format (.3gp)

The value of Intel NetStructure Host Media Processing Software lies largely in its reduced Total Cost of Ownership (TCO), often 50% less than traditional DSP-based solutions [12]. Because HMP software does not require specialized hardware, it is able to take advantage of economies of scale, resulting in lower hardware costs and some free performance improvements. As Intel develops better CPUs, HMP software performs better. It rides the Intel technology wave of silicon and software advancements: dual-core processors, HT Technology, Intel IPP, NICs, and VTune Performance Analyzer among others.

Because Intel NetStructure Host Media Processing Software runs on a general-purpose computer, development, deployment, and maintenance costs are reduced as well. Table 3 and Table 4 illustrate these points. Table 3 describes a hardware-based solution and a solution based on Intel NetStructure Host Media Processing Software. Table 4 analyzes the costs of the two solutions [12]:

Table 3: Hardware and Intel NetStructure Host Media Processing Software Solutions

Expense	Hardware-Based Detail	HMP Software-Based Detail
Basic Building Blocks	Intel NetStructure IPT2400 + Intel NetStructure DMV2400AC PCI	Resources: 240 Voice, 240 RTP, 60 CSP + server with Intel® Xeon® processor (3.2 GHz -- $3,500)
Development System	10% of deployed ports	5% due to NFR software pricing
Inventory	25% of volume	None
Shipping	Ship boards from distributor	Electronic distribution
Installation and Configuration	4 hrs @ $100 per hour	1 hr @ $100 per hour
Spares	30% of boards	Instant emergency license
Field Upgrade + 25% Capacity	Buy Intel NetStructure® DMIP301 boards	Add 60 Ports + additional server

Table 4: Analysis of both solutions

Expense	HW-Based Cost	HMP Software-Based Cost	Savings
Basic Building Blocks	$21,260	$11,300	$9,960
Development System	$2126	$650	$1,476
Inventory	$1,875	$0	$1,875
Shipping	$100	$0	$100
Installation and Configuration	$300	$100	$200
Spares	$2,250	$0	$2,250
Field Upgrade + 25% Capacity	$6,378	$5,456	$922
Total	$34,289	$17,506	$16,783 (49%)

® Intel and Xeon are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Intel APIs

Intel NetStructure Host Media Processing Software supports the R4 and Global Call APIs, which have been staples of computer telephony for over a decade. Both are high-level "C" APIs that make implementing media servers easier for the application developer. The APIs are the same APIs that run on Intel media processing technology boards, making the transition from board-based solutions to HMP software-based solutions nearly seamless. Connection control is internal if it is internal to the HMP software, and external if it connects the HMP software to an external entity. Global Call is the signaling and external connection control API, and R4 is the media control and internal connection control API.

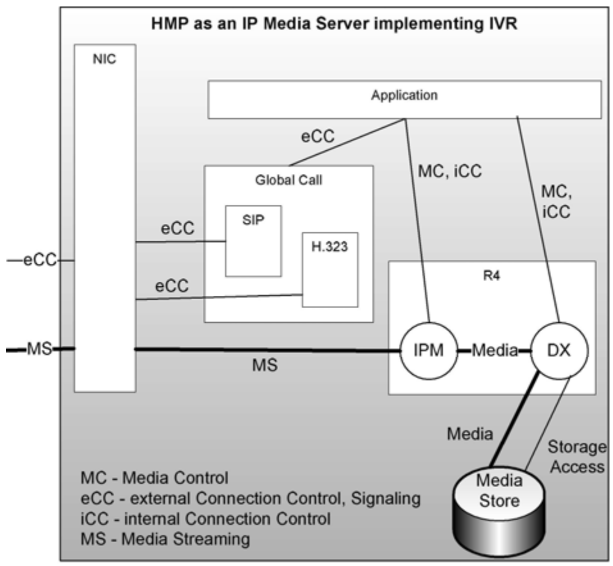

Figure 6: HMP software-based IP media server

Global Call abstracts the H.323 and SIP signaling stacks and can be configured to automatically establish IP calls or let the application manage the connections. After setting up calls with Global Call, an application performs the media control functionality through the R4 API. The R4 API is divided into several families [13-17]:

Table 5: R4 API families

API Support	
Voice processing	R4 voice (dx_)
Multimedia processing	R4 multimedia (mm_)
Connections	R4 routing (sc_ and dev_)
Conferencing	R4 conferencing (dcb_)
Fax	R4 fax (fx_)
Continuous speech processing	R4 EC (ec_)
IP media (QoS, etc.)	R4 IPML (ipm_)
Framework - Event reporting, device enumeration, and other related functionality	R4 SRL (sr_)

R4 abstracts media control functionalities into R4 devices. Example devices are the DX device, which performs audio play and record, the MM (multimedia) device, which performs multimedia play and record, and the IPM (IP media) device, which performs RTP functionality. Applications use these devices not only to perform actions, but also to receive events and make connections. All devices support duplex media streaming. Some of these functions are described in more detail below.

R4 includes two more abstractions that are worth an introduction: Run-Time Controls (RTCs) and termination conditions. RTCs allow applications to arm devices to respond to events without application interaction. For instance, the application can arm the voice device to increase volume on detection of a digit. In essence, RTCs are a means to meet real-time performance requirements.

A termination condition is a condition associated with an action, such as record, that should terminate upon occurrence of the condition. Setting up termination conditions is a common use of RTCs. A common termination condition is silence duration, which might trigger a voice device to terminate recording of a media stream.

R4 and Global Call both support a variety of programming models, broadly categorized as synchronous and asynchronous. Application developers are encouraged to use the asynchronous programming model for superior performance.

In the asynchronous programming model, the action initiated by a function call, such as playing, does not complete when the function returns. Instead, HMP software sends an asynchronous completion event to the application once the action has completed. By doing this, an application does not have to wait for one action to complete before starting another, and the application does not have to spawn a thread for every channel in the system.

A robust technique for asynchronous programming is to program an application state model. In this state model, events, including completion events, incite the state transitions. Hence, an application event handler checks the state of the device on which the event is received in order to determine the next action.

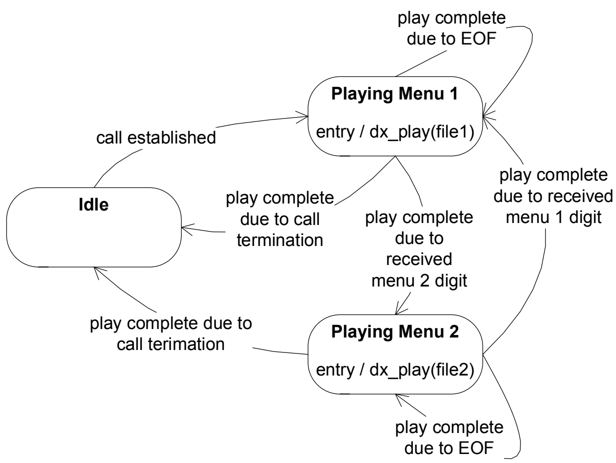

Figure 7: Simple IVR application state diagram

Figure 7 is an example of an Interactive Voice Response (IVR) application state diagram. The "entry/dx_play(...)" notation indicates that the application performs a dx_play() on entry into that state (more on dx_play() later). By using an application state model, the vast majority of the business rules for an application is built into the model, making the application not only robust but also flexible.

Building an HMP Software Application

This section applies our Intel NetStructure Host Media Processing Software and API knowledge towards building an IVR and Video Portal application. This section does not provide an exhaustive discussion or code listing, but does briefly address all of the key pieces. Finally, this section uses pseudo-code with function names identical to R4 and Global Call function names.

An IVR application has four main pieces: establishing a call, connecting an IPM device to a local voice device, connecting an IPM device to a remote endpoint, and controlling the media in the call through the voice device. We will tackle these pieces in that order. The application design is consistent with Figure 7.

Global Call provides the APIs for making and receiving calls [18].

```
gc_MakeCall() // to make a call
gc_WaitCall() // to wait for a call
```

After these APIs complete, via asynchronous completion events, a variety of other Global Call APIs are used to complete the external connection setup. Before we connect an IPM device to a remote endpoint, we need to connect it to a local voice device using an internal connection control.

```
dev_Connect(IPM_device, voice_device, duplex)
```

And eventually, we need to configure our IPM device:

```
ipm_StartMedia(IPM_device, audio codec, UDP_port,
        IP_address, ...)
```

The codec, UDP port, and IP address information are available from Global Call. At this point we use Global Call to inform the far side that the connection is up and running.

Now we are ready to perform IVR functionality, for example, play a file from the voice device.

```
dx_play(DX_device, file);
```

That completes the critical steps for building an IVR application. Each step above is probably three to five steps for most applications. A typical IVR application will typically perform multiple audio plays, depending upon input received from the user. Furthermore, the application needs to be constructed asynchronously, as described above, but the concepts and steps are intuitive.

To change our IVR application into a video portal, we need to add video functionality. In HMP software from Intel, video is handled through the multimedia libraries (mm_), which differ from the voice-only libraries (dx_). Therefore, we need to change the dx_ calls to mm_ calls, but we retain the same state machine. For example:

```
dev_Connect(IPM_device, DX_device, duplex) ⟶

dev_Connect(IPM_device, MM device, duplex)

dx_play(DX_device, file) ⟶

mm_Play(MM_device, file)
```

Most dx_ calls have analogues in the mm_ domain. The parameters are arranged a little differently, but the style is the same, so switching from dx_ to mm_ is a natural and necessary process, as the dx_ calls could not support video requirements.

In addition, we must update our Global Call code and ipm_StartMedia to include video stream configuration:

```
ipm_StartMedia(IPM_device, IPM_device, audio codec, video codec, UDP_ports, IP_address, ...)
```

Again, Global Call provides the codec and port information.

That completes the transformation of our application from IVR to Video Portal. As we have seen, Intel NetStructure Host Media Processing Software is a powerful tool for implementing media applications, including media servers. The HMP software reduces application development time and TCO. Now we look at some of the Intel technologies on which Intel NetStructure Host Media Processing Software is built.

INTEL ARCHITECTURE FOR SIGNAL PROCESSING APPLICATIONS

Intel offers an exceptional signal-processing platform for building next-generation media servers with the lowest total cost of ownership.

Intel® Integrated Performance Primitives

Intel IPP are a highly optimized suite of libraries for audio, video, imaging, cryptography, speech recognition, and signal processing functions [22]. To maximize performance, Intel IPP use advanced performance-tuning techniques such as pre-fetching and cache-blocking, avoiding data and trace-cache misses as well as branch mis-predictions. Intel IPP exploit instruction set architectures like Intel Wireless MMX™ technology, SIMD Extensions (SSE), and HT Technology.

The Intel IPP libraries can be linked to the application as dynamically loadable modules, which make the applications platform independent. The libraries automatically detect the underlying processor platform at run-time and execute the function implemented for that particular platform. Table 6 lists some of the Intel IPP that are related to Media Processing. Intel NetStructure Host Media Processing Software uses the Intel IPP for high performance. Figure 8 demonstrates the performance gain offered by IPP over Compiled C code [22]. In Figure 9 we show IPP encoder performance for H.263 profile 0, QCIF at 15 frames per second.

Table 6: Intel IPP

Media Processing Function	Intel IPP Available
Audio/Video Play and Record, Audio and Video Transcoding for multi-media connections over IP.	**Audio Codecs** – G.711, G.723.1, G.729, G.722, G.722.1, G.726, G.728, GSM-AMR, MP3, AAC, AC3. **Video Codecs** – H.263, H.264, MPEG4
AGC, Tone Detection, Tone Generation, VAD, Conferencing Summer	**Signal Processing and Vector Math Library** – Digital filtering (Adaptive, FIR, IIR), Fourier Transforms, Signal Generation
Speech Recognition	**Audio Processing** – Acoustic Echo Cancellation, Noise Reduction, VAD, Feature Extraction. **Speech Processing** – Pitch Detection, Speech Resampling
Echo Cancellation	G.168-2000 compliant Echo Canceller
Secure RTP	Symmetric Cryptography (DES, 3DES) Hash Algorithm Data Authentication (DES, 3DES)

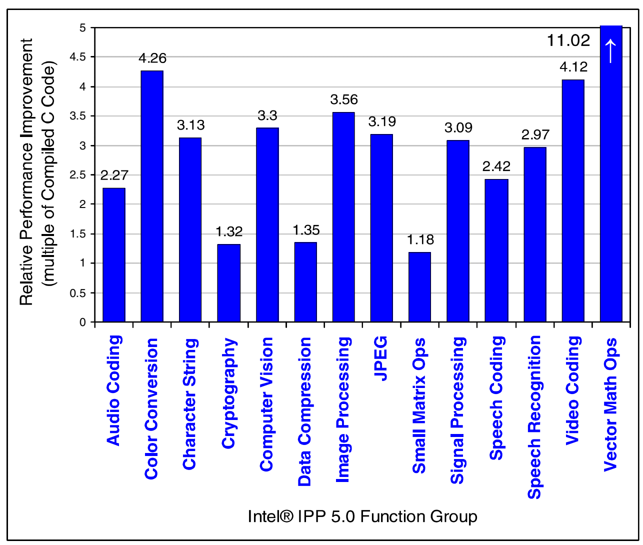

Figure 8: Intel IPP 5.0 average performance gain over compiled C code[3][22]

™ Intel Wireless MMX is a trademark or registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.

[3] All code running on a PC with an Intel® Xeon® processor supporting Hyper-Threading Technology, 3.6 GHz, 1 MByte L2 cache and 2 GB RAM using Microsoft Windows* XP.

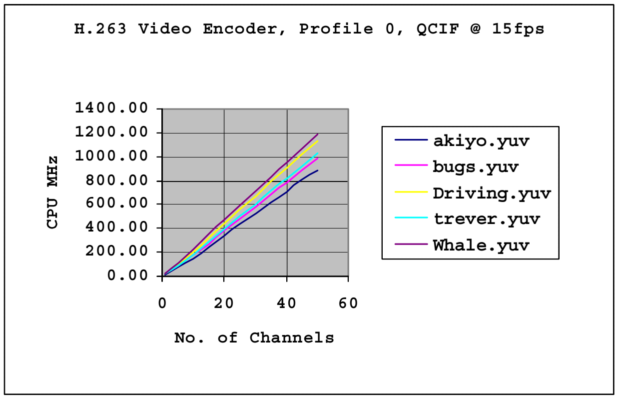

Figure 9: H.263 IPP encoder performance[4]

Figure 10 shows performance data for advanced video processing functions such as text overlay and tiling for video conferencing with video streams of QCIF @ 15fps.

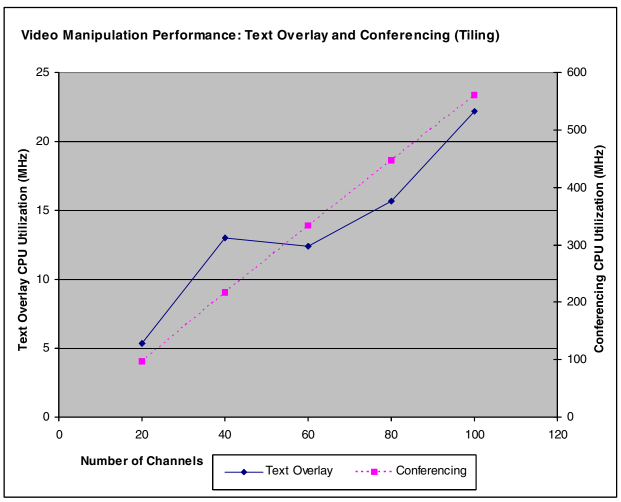

Figure 10: Video manipulation performance[4]

Processors and Chipsets

Intel offers a wide variety of processors and chipsets to address the desktop, server, and mobile market segments. Table 7 lists a sampling of the processors available today, addressing each of the market segments. These processors offer a combination of features and technologies such as HT Technology, dual-core processors, increased amounts of L2 and L3 cache, increased FSB speeds, Intel Extended Memory 64 Technology (Intel EM64T), Streaming SIMD Extensions 3 (SSE3), and multi-processor platforms (MP). The hardware platform combined with the software tools already mentioned makes Intel Architecture Processors a very compelling media processing hardware platform from both a functionality and a density standpoint. The impact of HT Technology, dual-core processors, multi-processors, cache sizes, Intel C++ Compiler, VTune Performance Analyzer, and Intel IPP on media processing is discussed in detail. All data presented in this paper have been generated on a specific platform configuration that may vary for each set of results. All numbers presented are representative of the platform configuration for only that given set of data.

Table 7: Sampling of Intel processors

	Pentium® Processor Extreme Edition 840[5]	**Pentium® 4 Processor Extreme Edition**	**Intel® Xeon® Processor Dual-Core**	**Intel® Xeon® Processor MP**
System Type	UP	UP	DP	MP
L2 Cache	2x1MB	512 KB	2x2MB	1MB
L3 Cache	N/A	2 MB	N/A	8 MB
Clock Speed	3.20 GHz	3.46 GHz	2.80 GHz	3.33 GHz
FSB (MHz)	800 MHz	1066 MHz	800 MHz	667 MHz
dual-core	✓		✓	
Intel® EM64T	✓		✓	✓
HT	✓	✓	✓	✓
Execute Disable Bit	✓		✓	✓
SSE3			✓	
EIST			✓	
DBS			✓	✓
Chipset	Intel® 955X Express	Intel® 925XE	Intel® E7520	Intel® E8500
Memory Type	Dual-Channel DDR2	Dual Channel DDR2 400/533 (CL3)	Dual Channel DDR, DDR2	Quad Channel DDR, DDR2

[4] All code running on a PC with an Intel® Pentium® M processor supporting Hyper-Threading Technology, 2.1 GHz, 2 MByte L2 cache and 1 GB RAM using Microsoft Windows* XP.

[5] Intel processor numbers are not a measure of performance. Processor numbers differentiate features within each processor family, not across different processor families. See http://www.intel.com/products/processor_number for details.

Table 8: Overview of Intel processor technologies

System Type	Number of processor sockets in a platform or computer system. Uni-processor means that only one processor can be present. Dual-processor systems allow for up to two processors, and multi-processor systems allow for more than two processors.
L2 Cache	Ultra-fast memory that buffers information being transferred between the processor and the slower RAM in an attempt to speed these types of transfers.
L3 Cache	Size of 3rd-level cache, typically larger than L2. L3 Cache is ultra-fast memory that buffers information being transferred between the processor and the slower RAM in an attempt to speed these types of transfers. Integrated L3 cache provides a faster path to large data sets stored in cache on the processor.
FSB (Front Side Bus)	Bus connecting the processor to the main external memory. The frequency shown in the table represents the operating frequency of the bus.
Intel® EM64T	Intel Extended Memory 64-bit technology enables 64-bit computing.
Execute Disable Bit	Allows the processor to classify areas in memory where application code can execute and where it cannot.
SSE3	Internet Streaming SIMD (Single Instruction Multiple Data) Extensions are instructions that reduce the overall number of instructions required to execute a particular program task. 3 refers to the 3rd iteration of these enhanced instructions.
EIST	Enhanced Intel SpeedStep® technology enables a system to dynamically adjust processor voltage and core frequency.
DBS	Demand-Based Switching uses EIST to dynamically lower the processor voltage and core frequency based on processor utilization.

Hyper-Threading Technology

HT Technology boosts computing performance by enabling a single processor to function as two "virtual" processors by executing two threads in parallel, allowing software to multi-task more effectively [23]. It provides thread-level parallelism on a single processor, resulting in more efficient use of processor resources, higher processing throughput, and improved performance of multi-threaded applications. Multi-threaded software divides its workloads into processes and threads that can be independently scheduled and dispatched by the operating system. In a multiprocessor system, those threads execute on different processors, whereas in a single processor that is HT Technology enabled, the threads execute in parallel on a single processor. HT Technology uses the idle cycles in a processor core such as stalls due to memory access, to enable parallel execution of another thread such as arithmetic computation that utilizes internal registers.

Prior to HT Technology, media-processing functions such as playing audio (disk or memory I/O intensive) and decoding a g.729a VoIP packet would complete in order of priority. If the thread executing audio play happened to be higher priority, the g.729a decode thread would not be able to take full advantage of the stalled processor cycles in the audio play thread due to memory or disk I/O. With HT Technology enabled, the g.729a decode thread can execute in parallel during the stalled cycles by doing any arithmetic operations that do not require memory I/O.

Figure 11 and Figure 12 show the performance due to HT when compared with single- and dual-processor systems for a specific media processing application: the SIP-to-SIP connection is configured for G.711 and G.729a RTP connection. For a G.711 RTP connection, HT Technology results in close to a 50% reduction in CPU utilization–equivalent to executing on a dual-processor configuration. For a G.729a RTP connection, HT Technology results in a 15-20% reduction in CPU utilization, whereas a dual-processor configuration without HT Technology results in a 30-40% reduction in CPU utilization. In both test application configurations, the platform comes close to achieving the theoretical maximum of four times the performance increase when compared to a dual-processor which is two times the performance speed, with HT Technology-enabled (x2) platform with a single-processor platform without HT Technology. The difference can be attributed to more frequent conflicts between the various channel threads using shared resources such as the execution engine, external memory access, and cache. If there is contention in usage of these resources by two different threads, the operations are serialized. A G.711 encode/decode algorithm has a small code and data footprint that will most likely not result in any cache misses. The G.729 algorithm has a much larger data footprint and data structure per channel and there is likely to be more cache thrashing and contention for external memory access.

® Intel SpeedStep is a trademark or registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.

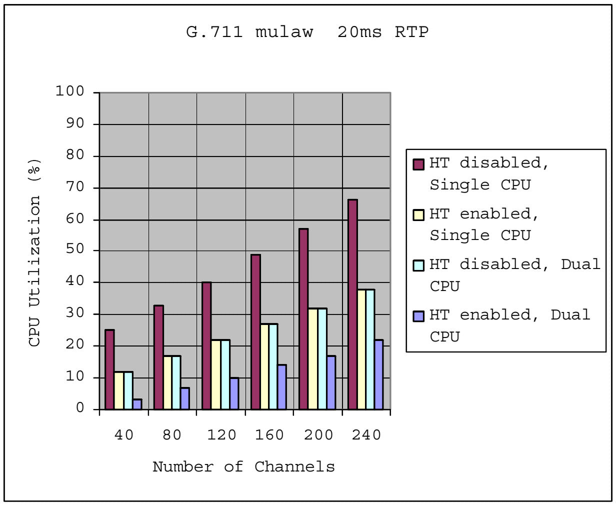

Figure 11: HT technology and dual processor performance G.711[6]

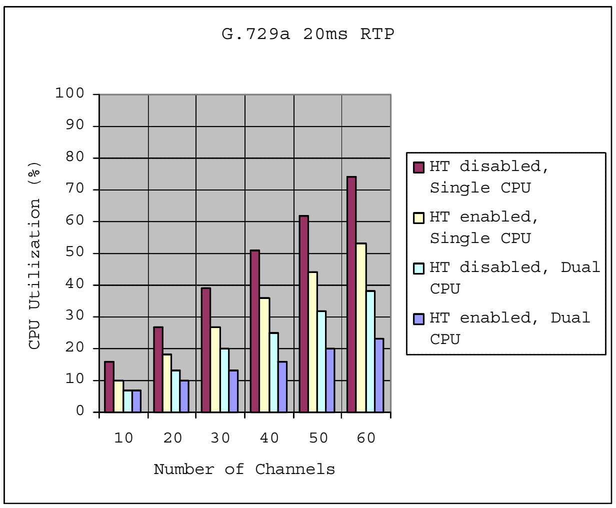

Figure 12: HT technology and dual processor performance G.729[6]

Dual-Core Processors

Dual-core processors provide increased computing performance by combining two full processing cores into a single processor. These processors are well suited to multi-tasking environments because there are two complete execution cores instead of one, each with an independent interface to the FSB. Unlike HT Technology, dual-core processors offer complete parallelism of threads through independent execution units with no contention between threads in accessing resources such as the execution engine, FSB, and cache. Dual-core processors also open up a multitude of opportunities for executing completely independent tasks on each processor such as gaming on one core while running a virus scan on the other. The dual-processor results in Figure 12 indicate the potential performance boost due to dual-core processors.

Cache–Intel Architecture Processors

Cache Intel Architecture processors are offering increasing amounts of L2 and in some cases L3 cache, with significant potential benefits for cache-friendly applications. Media-processing applications benefit from increasing levels of cache as CPU-intensive signal-processing algorithms execute more quickly, with fewer processor stalls associated with fetching instructions and data from memory. A typical signal-processing chain for an audio play application to a SIP endpoint involves fetching a block from memory or file, parsing the block for header and raw data, feeding the raw data into a decoder, and adjusting gain by a gain block. The output of the gain block is then fed into an encoder, followed by packetization into RTP packets. A multi-channel voice mail server would repetitively perform the signal processing functions associated with audio play with the instructions executing out of cache. Data for each channel are fetched only at the start of the processing with subsequent blocks accessing the data out of cache. Large amounts of L2 and L3 cache ensure that there are fewer cache misses during the entire signal-processing chain. Figure 13 and Figure 14 show the performance boost (reduction of 12-15% in CPU utilization) due to L3 cache for an audio play application.

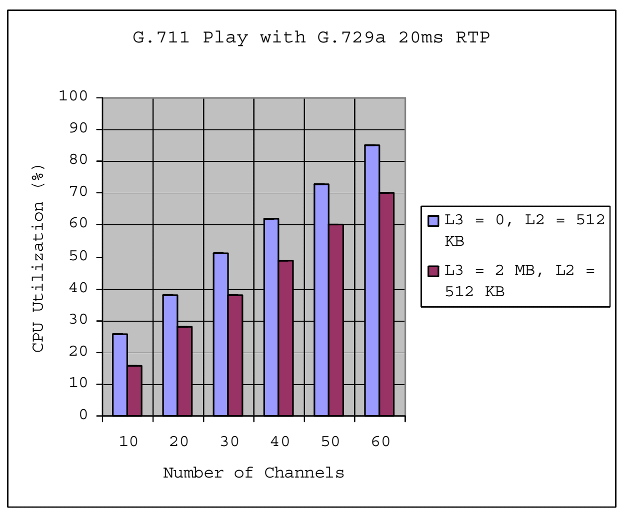

Figure 13: Cache performance G.729a[7]

[6] All code running on a PC with dual Intel® Xeon® processors supporting Hyper-Threading Technology, 3.2 GHz, 512 KByte L2 cache and 1 GB RAM using RedHat* Enterprise Linux.

[7] All code running on a PC with an Intel® Pentium® 4 processor supporting Hyper-Threading Technology (disabled), 3.2 GHz, 512 KByte L2 cache, 2 MByte L3 cache and 1GB RAM using RedHat* Enterprise Linux.

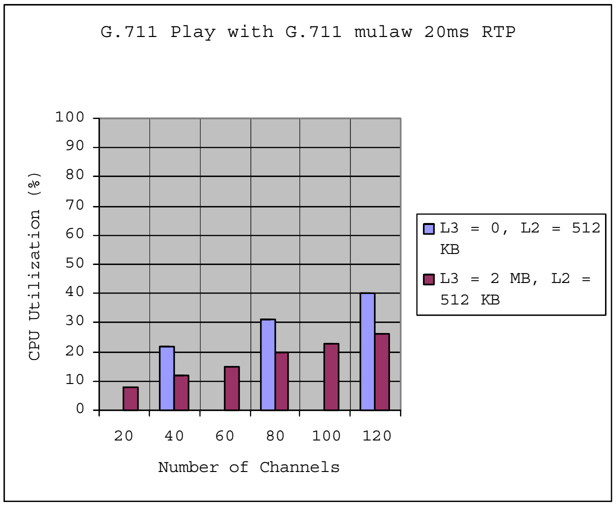

Figure 14: Cache performance G.711[7]

INTEL DEVELOPMENT ENVIRONMENT

Intel VTune Performance Analyzer

Intel VTune Performance Analyzer is a powerful tool for analyzing the performance of an application. VTune analyzer has two main modes, sampling and call graph mode [20].

Sampling mode is a non-intrusive way to profile the entire system. In sampling mode, statistics are rolled up to the application level. The user can then delve down to the function level. The user can double-click on a function and get instruction-level statistics as annotations in the source files [20].

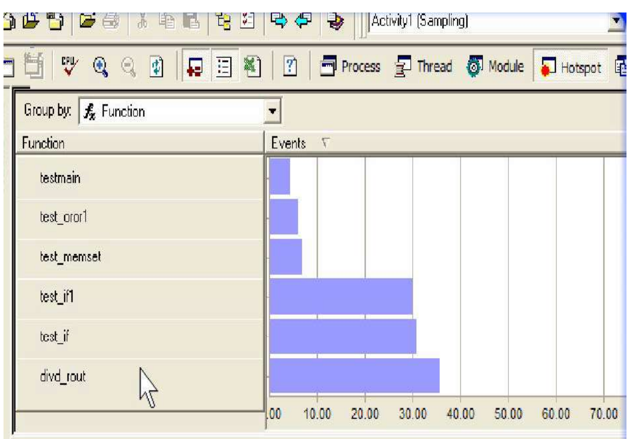

Figure 15: VTune Performance Analyzer sampling mode output [20]

Call graph mode is more invasive and slows the program under test. It does, however, enable a graphic view of calling sequences of the different procedures within the program. It further identifies critical paths in the program. Statistics are provided for each procedure, including "wait time," the amount of time the function spends waiting for an event to occur.

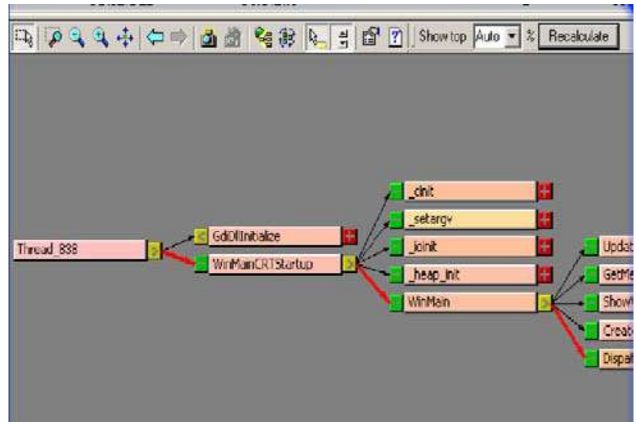

Figure 16: VTune Performance Analyzer call graph mode output [20]

Beyond CPU utilization and wait time, VTune Performance Analyzer collects important statistics such as cache misses, clock ticks per instruction, and thread utilization.

Under Windows, another tool, VTune Performance Analyzer Tuning Assistant, makes recommendations on improving the performance of certain functions.

Use of VTune Performance Analyzer can be thought of as a three-step process: set up, find an area to improve, and improve. Finding an area to improve, given VTune analyzer's intuitive user interface, usually takes no more than an hour.

VTune analyzer's output is function-centric. In some cases, it is nice to abstract function-level results into higher-level module results. A parsing utility can easily be used to aggregate function-level statistics into module-level statistics.

VTune Performance Analyzer should not be used as a primary tool for the measurement of performance. VTune analyzer's role is in improving performance. However, VTune analyzer can be useful for understanding the performance characteristics of one release over another. For instance, suppose we upgrade a software release from A to B. If, as a matter of course, we measure A and B with VTune analyzer, we can compare their performance characteristics. Are we seeing the same relative performance of our code? Has one module significantly decreased in performance? Does this make sense? While use of this A-B sanity check is not strictly necessary, it can help rapidly identify probable design or coding inefficiencies.

VTune Performance Analyzer is an important tool for getting maximum performance out of the Intel

Architecture. Intel NetStructure Host Media Processing Software, in particular, benefits from regular inspection with VTune analyzer.

Intel C++ Compiler

The Intel C++ compiler is one of the software development tools available to accelerate software performance on Intel platforms. It is available for a number of different operating systems, including Windows and Linux. The Intel C++ Compiler 9.0 for Linux provides outstanding application performance for software running on Intel processors [25]. It includes advanced optimization features such as full support of multi-core processors with capabilities including Auto-Parallelization, Optimized floating point instruction throughput, Interprocedural Optimization (IPO), Profile-Guided Optimization (PGO), and Data prefetching**.** It includes a Compiler Code-Coverage tool and a Compiler Test-Prioritization tool. Optimizations specific to the IA-32 architecture are provided, such as full support for Streaming SIMD Extensions 3 (SSE3)**,** Automatic vectorizer, and Processor Dispatch, as well as support for Intel Extended Memory 64 Technology. More details are available in "Intel C++ Compiler for Linux" [25]. The compiler also includes an enhanced debugger that allows debugging of optimized code, as well as support for stack frame runtime error checking to help reduce buffer overrun security exploits. Various case studies are presented in "Intel Software Tools Case Studies" [27], highlighting performance improvements due to Intel software tools. Table 9 highlights the performance improvement provided by Intel Compiler and IPP in one particular case study–H.263 Video Encoding [26].

Table 9: H.263 Encode Performance Improvement due to Intel Compiler and IPP [26]

ImageCom PC Encoder Configuration		Intel Pentium 4 processor-based system	
Intel IPP	Intel Compiler	Encode Time for 80-sec clip	Percentage Improvement
		123 sec	0%
Y		84 sec	32%
	Y	69 sec	44%
Y	Y	57 sec	54%

WHERE WE GO FROM HERE

While this paper has primarily focused on media servers from the point of view of traditional telecommunication applications, media servers will play an important role in the next-generation digital home as well. As broadband to the home becomes commonplace, we will see the emergence of new and exciting applications that will have a big impact on our personal lives. We are already seeing the emergence of TV over IP (IPTV) promising to change the way we access public content such as broadcast and on-demand programming. Providers will be able to personalize what their customers receive; not only providing them with content that interests them but also content when they want it.

Going beyond public content, media services will be used in new and innovative ways to enrich our personal lives. For example, media servers may be used to share private content such as personal photos and video with friends and families, and to do so securely. Residential media servers may be used for home security and health monitoring as well as entertainment.

Intel technology will continue to advance to meet the demands of these future applications. With leadership technologies such as multi-core, I/O acceleration, virtualization, power management and security, to name just a few, Intel promises to continue to provide platforms that exploit the next generations of platform technologies to their fullest.

CONCLUSION

A media service network is a network through which a wide range of media services are provided by application programs controlling media resource functions on a media server. Intel offers a wide spectrum of standards-based technologies that facilitate the building of flexible, high-performance, low-cost media servers that can be deployed in the circuit-switched, pure-IP and next-generation converged networks.

We have seen that Intel NetStructure Host Media Processing Software is a feature-rich platform, requiring no special-purpose hardware, on which a wide variety of media applications can be developed for the next generation of IP-based networks. Utilizing industry standards and the popular Windows and Linux operating systems, HMP software-based solutions can be built with a significantly lower total cost of ownership and shorter time to market than proprietary hardware-based solutions. We walked through a simple example of how the high-level abstractions provided by the APIs in Intel NetStructure Host Media Processing Software make application development straightforward.

We demonstrated how low-level building blocks like Intel IPP make Intel Architecture processors capable of functionality previously reserved for special-purpose DSPs. The popular Intel Architecture processors provide a cost-effective infrastructure for media signal processing that is high in performance as well as economical.

Tools such as the VTune Performance Analyzer, the Intel C++ compiler, and Intel IPP enable software technologies, including Intel NetStructure Host Media Processing Software, to get the absolute highest performance from the hardware platform.

Leading the advances in technologies such as multi-core, I/O acceleration, virtualization, power management, and security, Intel promises to continue providing technologies and platforms that enable cutting-edge media services.

PERFORMANCE TESTING

All testing was performed internally by Intel.

Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, visit http://www.intel.com/performance/resources/limits.htm.

ACKNOWLEDGMENTS

We thank all of our colleagues who reviewed and commented on this work. We especially acknowledge Asha Abraham, Rangarajan R. Calyanakoti, Ashok Kondru, Steve Masson, Bhaskar Rao, Daniel Rhee, and Glen Shires, without whose insightful technical contributions we could not have produced this paper.

REFERENCES

[1] "Introduction to CCITT Signaling System No. 7," *ITU-T Q.700* (03/93), March 1993.

[2] "ISDN user-network interface layer 3 specification for basic call control," *ITU-T Q.931* (05/98), May 1998.

[3] "Packet-based multimedia communications systems," *ITU-T H.323* (11/00), Nov. 2000.

[4] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler, "SIP: Session Initiation Protocol," *IETF RFC 3261*, June 2002.

[5] Schulzrinne, H., Casner, S., Frederick, R. and V. Jacobson,"RTP: A Transport Protocol for Real-Time Applications," *IETF RFC* 1889, Jan. 1996.

[6] "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; IP Multimedia Subsystem (IMS); Stage 2 (Release 5)," *3GPP TS 23.228* V5.5.0, 2002-06.

[7] M. Arango, A. Dugan, I. Elliott, C. Huitema, S. Pickett, "Media Gateway Control Protocol (MGCP)Version 1.0," *IETF RFC 2705*, Oct. 1999.

[8] H. Schulzrinne, S. Petrack, "RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals," *IETF RFC 2833*, May 2000.

[9] "Voice Extensible Markup Language (VoiceXML) 2.1," *W3C Working Draft 28*, July 2004.

[10] Melanchuk, T., Sharratt, G., "Media Sessions Markup Language (MSML)," *draft-melanchuk-sipping-msml-03*, Aug. 15, 2004.

[11] Melanchuk, T., Sharratt, G., "Media Objects Markup Language (MOML)," *draft-melanchuk-sipping-moml-03,* Aug. 15, 2004.

[12] Intel Corporation, "A Comparison of Intel NetStructure® Host Media Processing Software and Traditional Solutions," http://www.intel.com/network/csp/applnots/9331an.pdf, 2005.

[13] Intel Corporation, "Intel NetStructure® Host Media Processing Software Release 2.0 for Windows," http://www.intel.com/design/telecom/prodbref/8762.pdf, 2005.

[14] Intel Corporation, "Multimedia API Library Reference," http://www.intel.com/design/network/manuals/telecom/hmp15lin/pdffiles/multimedia_api_v1.pdf, Sep. 2005.

[15] Intel Corporation, "Voice API for Host Media Processing Library Reference," http://www.intel.com/design/network/manuals/telecom/hmp15lin/pdffiles/voice_api_hmp_v2.pdf, Apr. 2005.

[16] Intel Corporation, "IP Media Library API for Host Media Processing Library Reference," http://www.intel.com/design/network/manuals/telecom/hmp15lin/pdffiles/ip_media_api_hmp_v5.pdf, Aug. 2005.

[17] Intel Corporation, "Device Management API for Windows* and Linux* Operating Systems Library Reference," http://www.intel.com/design/network/manuals/telecom/hmp15lin/pdffiles/device_mgmt_api_v3.pdf, Aug. 2005.

[18] Intel Corporation, *"*Global Call API for Host Media Processing Library Reference,*"* http://www.intel.com/design/network/manuals/telecom/hmp15lin/pdffiles/globalcall_api_hmp_v3.pdf, Aug. 2005.

[19] Intel Corporation, "Intel NetStructure® Host Media Processing Software*,"* http://www.intel.com/network/csp/promo/9403web.htm, 2005.

[20] Intel Corporation, "Intel® VTune™ Performance Analyzer Flash Demo*,"* http://download.intel.com/software/products/vtune/downloads/vtune_v7_b42.exe, 2003.

[21] Intel Corporation, "Intel® VTune™ Performance Analyzers*,"* http://www.intel.com/cd/software/products/asmo-na/eng/vtune/index.htm, 2005.

[22] Intel Corporation, "Intel® Integrated Performance Primitives 5.0*,"* http://www.intel.com/cd/software/products/asmo-na/eng/perflib/ipp/index.htm, 2005.

[23] Intel Corporation, "Hyper Threading Overview*,"* http://www.intel.com/business/bss/products/hyperthreading/overview.htm, 2005.

[24] Intel Corporation, "Intel® Integrated Performance Primitives 5.0 performance data*"* http://www.intel.com/cd/software/products/asmo-na/eng/perflib/ipp/238657.htm, 2005.

[25] Intel Corporation, "Intel® C++ Compiler for Linux**"* http://www.intel.com/cd/software/products/asmo-na/eng/compilers/clin/219622.htm, 2005.

[26] ImageCom Inc, "A short cut to success*,"* http://cache-www.intel.com/cd/00/00/21/92/219258_imagecom.pdf, 2005.

[27] Intel Corporation, "*Intel® Software Tools Case Studies,"* http://www.intel.com/cd/software/products/asmo-na/eng/casestudies/index.htm, 2005.

[28] Intel Corporation, *"Intel NetStructure® Host Media Processing Software Release 1.5 for Linux*,*"* http://www.intel.com/design/telecom/prodbref/9323.pdf, 2005.

AUTHORS' BIOGRAPHIES

Joseph Grecco is a system architect in DEG Architecture and Planning at Intel, Parsippany, NJ. He has 20 years of experience in the communications industry with over 10 years involved in the architecture and design of media servers, middleware, and APIs. He holds a B.S. degree in Computer Information Science from the New Jersey Institute of Technology. Joe can be reached by e-mail at joe.grecco at intel.com.

Mark E. Mize is a system architect of media systems at Intel in Parsippany, NJ. He has a B.S. degree in Mathematics and a B.S. degree in Computer Science from the University of Texas at Austin. Mark has worked on IP protocols and media system design for the last eight years. Mark can be reached by e-mail at mark.mize at intel.com.

Ranjan Singh received his B.Eng. degree from Victoria University of Technology in Melbourne, Australia. Since joining Intel in 1999, he has worked on the digital signal processing components within Intel's NetStructure products. His current interests include architecture of next-generation media server products for enabling video and other emerging applications. Ranjan can be reached by e-mail at ranjan.singh at intel.com.

New Uses, Proposed Standards, and Emergent Device Classes for Digital Home Communications

Mark Walker, Digital Home Group, Intel Corporation
Jim Edwards, Digital Home Group, Intel Corporation
Ylian Saint-Hilaire, Digital Home Group, Intel Corporation
Jeffrey Lee, Digital Home Group, Intel Corporation

Index words: VoIP, UPnP, DLNA, HCS, Digital Home

ABSTRACT

The exponential growth rate of Voice-Over-IP (VoIP) subscribers worldwide, increasing deployment of broadband Internet access to the home, and the steady increase in the number of installed home networks has created an opportunity to support new communication experiences in the digital home. The end result for users will be new IP-based devices, services, and capabilities that go far beyond what today's communication technologies can deliver. Enabling these new communication experiences will require a digital home communication architecture consisting of new standards and classes of communication devices.

In this paper, we summarize the results of a recent Intel Corporate Market Research (CMR) study on unmet consumer communication needs and highlight how the collected data suggests the need to develop a general, standards-based framework for digital home communications and the development of two specific new device classes. The two device classes and their uses are presented in detail. The requirements for the proposed digital communications framework are compared with the existing Digital Living Network Alliance (DLNA) framework for digital home entertainment [1]. We conclude with an overview of similar architectural components that will be needed to establish the digital home communications framework.

INTRODUCTION

Voice-over-IP (VoIP) is a term used to denote voice communications over the Internet. Current VoIP users generally fall into the early adopter category; that is, a small but rapidly growing group of consumers willing to tolerate less than stellar audio quality in exchange for features like very low-cost long-distance access to any other Internet-accessible VoIP phone in the world [2]. The evolution of VoIP from a computer hobbyist technology into a solution for mass-market communications is predicted to accelerate as major telco and cable providers enter the market, adding capacity and improving quality [2]. The recently developed 3rd Generation Partnership Project (3GPP) IP Multimedia Subsystem (IMS) recommendations for Session Initiation Protocol-Session Description Protocol (SIP-SDP)-based call control [3-4] have now established the required foundation for interoperability between IP communications providers. Finally, a number of Internet Engineering Task Force (IETF) signaling and network protocol standard recommendations that specify the operation of VoIP devices [5-8] have been undergoing a process of consolidation and profiling within the industry, greatly increasing assurances of phone-to-provider interoperability.

We first summarize the results of a recent CMR study on unmet consumer needs for communications. We describe two new digital communication device classes suggested by the CMR data: a Home Communications Server (HCS) and a Digital Communications Adaptor (DCA). The HCS is similar in function to enterprise digital private branch exchange (PBX) systems, but is tailored specifically for home use and provides services exposed on the home network via UPnP* technology [9]. The HCS fulfills unmet consumer needs by providing a single entity for managing all home IP communications in a manner that abstracts the differences between different forms of communications such as voice, email, and instant messaging. It empowers consumers with a high degree of control over family communications by enabling complete personalization of HCS behavior

* Other names and brands may be claimed as the property of others.

through UPnP standard interfaces. The DCA adapts legacy consumer electronics (CE) equipment such as the TV for use as a communications device on the home network. DCA devices capture and render voice and optionally video, pictures, audio, text, and video for richer communications. Both new device classes require a UPnP technology standards framework for digital home communications, similar to that established by the DLNA for networked entertainment in the digital home.

Finally, it must be noted that in this paper we proceed from the assumption, as does much of the current technical literature on VoIP, that IP-based communications will gradually replace Public Switched Telephone Network (PSTN) communications over time, due to lower overall costs to the consumer and the ability to support richer communications such as video, instant messaging and picture sharing. While many devices deployed during the transition from PSTN to IP will support and interoperate with both systems, these devices are not considered here.

SUMMARY OF RECENT RESEARCH ON UNMET CONSUMER COMMUNICATION NEEDS

An Intel CMR study was recently conducted with the goal of defining an end-user-based framework for communications usage models within the digital home. The research team was composed of cross-organizational, cross-capability individuals from various Intel divisions with responsibility for the digital home market and technology. A total of 21 in-home interviews conducted in two phases took place over a three-month period in two North-American cities. Specific goals of this study were as follows:

- To develop an understanding of what drives the need for communications.
- To generate a first-level assessment of how technology is used in the communication model.
- To identify specific unmet consumer needs that technology can address.
- To translate unmet needs into framework, device and application requirements.

The key findings of this study were as follows:

1. Communication is increasingly taking place within what the team described as a "virtual home" environment. Virtual home is a term employed to highlight the fact that Americans now spend as much time outside their primary residence communicating and coordinating as they do inside it, irrespective of family size. Typical venues where families and individuals spend time include children's sports events and practices, churches, gyms, automobiles, etc.
2. Two key forces that drive the need for communication are "staying in touch" and "planning." Staying in touch means using communication to achieve a feeling of togetherness. Planning is the organization of family daily routines and dealing with the logistics of those routines. Families in the study reported significant amounts of time planning the day and communicating those plans to family members.
3. Consumers are using an increasingly diverse and non-interoperable set of communication channels such as cell phones, pagers, e-mail, Instant Messages (IM), and landline phones, The diversity of these channels has created "islands" of multiple in-boxes for messages, contact lists, calendars, etc. Consumers reported spending significant amounts of time manually bridging information and content between channel applications.
4. Four general categories of unmet communication needs were identified in the collected data: Continuity, Management, Protection, and Closeness. The term *continuity* denotes activities such as successfully juggling diverse channel types and minimizing the amount of personal attention required to distinguish between *wanted* (friends, family) and *unwanted* communication (e.g., telemarketers), and making it easy to receive wanted communications. *Managing* refers to the desire to have always available appropriate contact information, especially for contacts deemed important by the consumer. Appropriate contact information includes the communication means preferred (cell phone, land-line phone, IM) by the person being contacted. *Closeness* is the need to use technology to maximize the sense of togetherness when family members and friends are apart. Finally, *protection* refers to the ability to easily block unwanted communications.

CMR Data Analysis and Requirements

Analysis of and requirements for solutions emerging from the CMR study data is enumerated in this section. The results serve to help guide the definition of technologies to address unmet communication needs. Highlights of the analysis and required solutions are as follows:

- While the traditional concept of home is being supplanted by the more recent concept of the virtual home, the data suggests that the physical home may provide a base for deploying technology solutions that combine personal communication devices in a manner that among other benefits, affords users a feeling of home while they are away from home. Home-based solutions will never replace away from home solutions like cell phones, but cell phones may leverage emerging VoIP standards to integrate with in-home solutions and deliver the virtual home user experience.

- A user-transparent, home framework for connecting diverse constellations of communication devices to each other and for sharing information may help address the problem of bridging disparate communication channels. The framework may establish at least the perception, through the connectedness of devices, of a centralized entity for managing all forms of family communications including calls, messages, alerts, calendars, etc. High connectedness permits tight integration with task management applications, such as calendar.
- Employing proven networking technologies and standards as in [1], e.g., Ethernet, 802.11, IPv4, HTTP, and UPnP, for maximum availability and/or instant access to the framework may address the desire for control and management.
- Providing access to and use of the framework from outside the home supports the notion of virtual home.
- Enabling full user personalization and control of wanted and unwanted contacts, personalized policies for call and message handling, etc. address the consumer desire for communication protection. Using the framework to abstract calls, IMs, etc. allows personalization of control to be universally applied regardless of the communication type.
- Richer communications experiences such as videoconferencing or sharing pictures of the family while talking help address the need to achieve a greater sense of closeness. Applications of this type can maximize the use of technology to achieve a greater sense of closeness by allowing users for example, to see each other while speaking.

NEW PROPOSED DIGITAL HOME COMMUNICATIONS FRAMEWORK AND DEVICE CLASSES

The conclusion that the physical home may provide a base for deploying communication technology solutions suggests a need for a home framework that facilitates the integration of in-home personal communication devices and extends personalized communication services to devices used outside the home.

The DLNA framework for network-connected entertainment devices (summarized below) is proposed as a model for developing a digital communications framework. In this section, we focus on new communication device classes that utilize the proposed home communications network to help mitigate the unmet consumer needs listed above. The first device proposed is the concept of a Home Communications Server (HCS). An HCS is similar in essential function to an enterprise PBX, but tailored specifically to home use models. The second category of devices represents client devices that depend on the presence of the HCS on the home network. Digital Communication Adaptors (DCAs) fall into this category and are detailed below. DCAs allow legacy television sets to be used for richer communications, such as videoconferencing. Other categories of client devices (not discussed here) that derive value from the HCS but are not necessarily dependent on it include cell phones and VoIP wireless handsets.

Requirements of the Home Communications Server

General requirements for HCS use are as follows:

1. HCS implementations must support the desire for high ease of use implicit in the unmet needs by allowing any newly purchased or “guest” communications device to transparently register and use the home VoIP service provider for both inbound and outbound calls. This basic HCS connection process should be as similar as possible to the current industry standards for connecting VoIP clients to a service provider. Outbound calls are placed as requests to the HCS.
2. HCS implementations must support the *Managing* need by allowing any net-connected device to browse listings of contacts. Contacts entries must be in a form to allow communications-capable clients to transparently make outbound calls using the channel preference of the person being contacted. HCS implementations must additionally support easy editing/updating of contact listings by authorized client devices and applications.
3. An HCS must also allow authorized applications and client devices to set incoming call policies for call routing, call blocking/blacklisting, call forwarding to voicemail, client ring order, client ring tone, and external call forwarding–all based on whose calling and other criteria, including time/date. HCS implementations must allow the same policies

to also be set for incoming IMs and e-mails, effectively eliminating the differences between calls and messages in the task of setting home policies.

In addition to the requirements for meeting unmet needs, two additional HCS requirements for enabling compelling new user experiences are:

1. Authorized client devices and applications must be able to actively monitor all extensions (including who they are connected to) and to set up specific groups of clients for conference calls on the fly.
2. HCS implementations must be easily integrated with other productivity and planning applications like calendar.

HCS Architecture

Figure 1 shows the HCS and other devices connected to the home network. The HCS is connected to the Internet through the router and may be bound to a specific VoIP service provider. The HCS additionally employs SIP-standard telephony and signaling [5-8] for connections to the service provider and for receiving/acknowledging incoming calls.

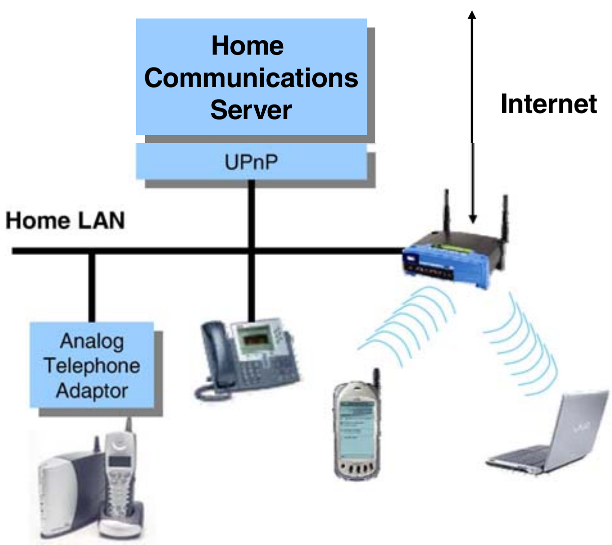

Figure 1: HCS on the Digital Home Communications Network

On the LAN side, the HCS exposes UPnP technology service interfaces to other networked devices that have UPnP control points. IP communication devices utilize their control points in the UPnP-standardized scheme to discover the HCS and obtain the IP addresses of the device and its service interfaces [9].

HCS services are grouped into four functional categories.

1. *Basic communication services*: After obtaining HCS IP addresses, IP communication client devices register with the HCS. Registered client information includes current IP address, friendly device identification, and audio capability profile. Audio capabilities are enumerated in the form of SDP parameter lists [10], that include the audio codecs supported by the device, ordered according to device preference.

   Once registered, IP communication client devices utilize the SIP proxy functionality of the HCS for placing outbound calls. SIP client devices literally call the HCS, which then forwards the call to the service provider. Calls inbound to the HCS from the service provider are routed to registered client IP addresses per call routing policies.

2. *Extended communication services:* Extended communication services allow the HCS to be managed and configured by any authorized client device. Management functions include the following:

   a. *Browse/search/edit/update contact listings:* These functions allow any device to view and edit XML-standard form contact listings. Contact metadata contains detailed technical information that enables among other attributes user selection of, the right "channel" when a user initiates an outbound call. Communication clients may first browse for a specific contact and include the contact listing reference in the outbound call request. Listings can also include recently dialed/received contacts. HCS users manage contact listings using standardized XML as the exchange format. This activity may be made even more user-friendly by supporting "1-button" contact list synchronization, i.e., a communications device with a native contact list can instantly synchronize with the HCS on its personalized contact subset.

   b. *Manage message archives*: This allows consumers to manage combined views of HCS voice, text message, and e-mail listings, as well as forward, delete, and save messages.

   c. *Manage home communication policies:* This allows users to manage XML-standardized HCS policies for these actions:

      - Routing of all incoming calls/messages/mails based on caller/sender ID to specific client extensions based on device ID.
      - Calls/messages/mail blocking based on time of day and caller/sender ID.

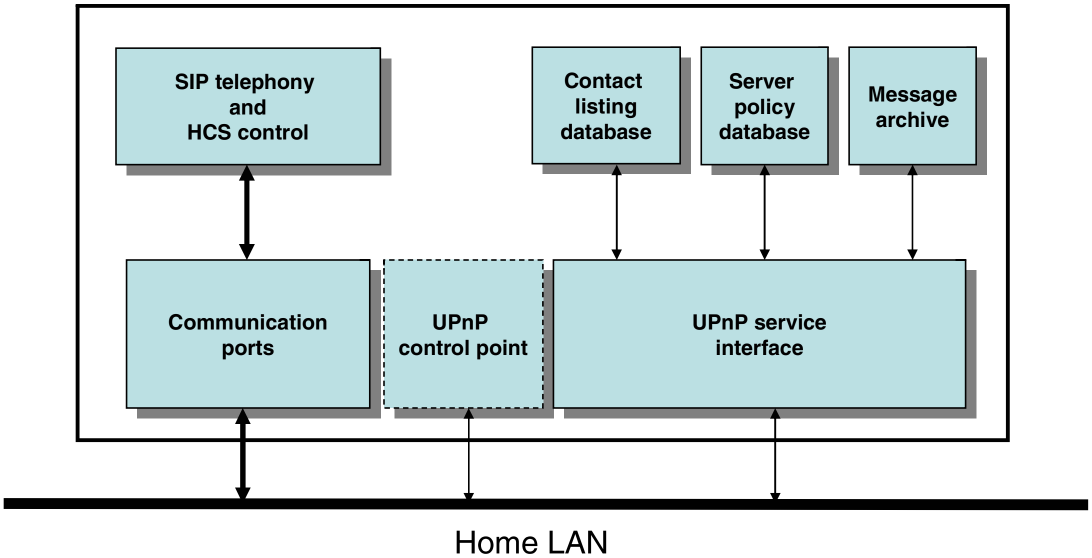

Figure 2: Functional blocks composing the HCS

- Call/message/mail priority: caller/sender ID-based priorities for incoming communications and device ID-based priorities for outbound calls.
- Caller/sender ID call/message/mail forwarding to voicemail.
- Client extension ring order.
- Caller/sender ID-based client extension ring/alert tone.
- Caller/sender ID-based external call/message/mail forwarding.

3. *Remote access***:** Family members can log into the HCS when outside the home and register their current IP addresses. The HCS then forwards inbound calls, messages, and mail to externally located users based on its forwarding policies. Using current standards for Internet security [11], users are additionally able to use HCS extended communication services, including the management functions described above.

4. *Communication notices/alerts:* The HCS may optionally include functionality for discovering digital media rendering devices (DTVs, digital stereos, etc.) connected to the home network, for the purpose of forwarding notifications of incoming calls and other alerts. In the next section, we detail a new device type that takes full advantage of the potential to merge communications and rich media.

Figure 2 shows a generalization of the functional blocks that compose the HCS. The contact listing database, server policy database and message archive are exposed directly through the UPnP service interface. The UPnP service interface also exposes communication ports available for outbound calls. The SIP telephony and HCS control module receives inbound calls, messages etc., and routes them according to policy to registered client devices. The optional control point can be used to discover and utilize other UPnP devices, including display devices.

Continuity, Management, and Protection

The HCS fulfills unmet consumer needs summarized above by providing a single entity for managing all home IP communications in a manner that abstracts the differences between different forms of communications such as voice, e-mail, and IM. It empowers consumers with a high degree of control over family communications by enabling complete customization of HCS behavior through standard UPnP interfaces.

Additional HCS Benefits to the Consumer

In addition to addressing the unmet consumer communication needs above, two additional benefits for home IP communication users are realized with an installed HCS:

1. *Call-quality assurances***.** Using the HCS as the proxy mechanism for initiating outbound calls allows the HCS to accept/reject call requests based on detected available bandwidth–on both the LAN and the LAN/WAN boundary. A user receiving an acknowledgement from the HCS for the requested call is generally assured of the bandwidth necessary to maintain a good communications experience. This will become an especially critical capability as network-connected digital entertainment devices are

deployed and begin consuming bandwidth on home networks.

2. *Voice interaction*. Advanced HCS implementations may include capabilities such as voice recognition and Text-To-Speech (TTS). Voice recognition can be utilized to establish voice user interface for command and control functions such as voice dialing and voice browsing/searching of contacts. TTS capabilities can be employed for talking IMs and e-mails.

PCs as HCS Devices

The PC possesses both advantages and challenges for implementing HCS functionality. The major advantages include computational power for rendering rich UIs used for communications management, storage capacity for storing messaging and contact info, and the ability to provide voice recognition and TTS applications. Challenges include a perceived gap in the level of robustness that consumers associate with CE appliances. While PC architecture is improving and the PC industry is working to address availability and reliability, an HCS implementation may be perceived as a single point of failure in case of power outage, denial-of-service (DOS) attacks, software, or hardware failure. Consumers accustomed to the reliability of traditional PSTN telephones might have concerns. For these reasons it seems reasonable to consider a deployment model in which the HCS functionality is split across two platforms connected to the home network. Basic communication functions would be deployed on a CE appliance device, possibly possessing an uninterrupted power supply or battery backup. The remaining extended capabilities like HCS management, etc. would be supported by the PC. In this manner the PC would function to greatly enhance the convenience, manageability, and personalization of home IP communications, and would also be the focus of bridging to IP-connected entertainment devices without being in the critical path.

Digital Communications in the Family Room

The presence of both communication and digital entertainment devices connected to the home network provides an opportunity for supporting communication experiences that leverage the rich media features of entertainment devices. High-definition televisions have the potential to deliver full family (rather than personal), videoconferencing, picture and music-sharing communication experiences in the family room in a manner more compelling than previous generations of video telephony devices.

Use cases for videoconferencing include the notion of "virtual presence." For example, family members unable to be together for the holidays establish a high-definition videoconferencing session between household members in the family room. Such sessions are enhanced by spontaneously sharing photos and music. Another type of virtual presence use case is the concept of persistent communication channels that enable, for example, the monitoring of elderly or infirmed individuals. Having direct support for control of persistent videoconferencing from the digital television (DTV) remote can effectively provide closeness with friends and family members in the form of "channels," controlled and selected in the same way that cable entertainment is controlled and selected.

The Digital Communications Adaptor

DCAs are used for personal videoconferencing and are architecturally similar to Digital Media Adaptors (DMAs) in that they rely on a connection to a home media server and function principally to render digital streams onto legacy televisions. DCAs differ in that they add 2-way audio and optionally, 2-way video capture and media sharing capabilities to the basic DMA profile. The DCA depends on the HCS to respond and connect it to incoming calls. All external telephony, links to service provider, contact listings, etc. are the responsibility of the HCS. DCAs are the lowest cost implementation for personal videoconferencing. DCAs achieve low cost by off-loading functionality to the server (HCS) such as contact archives, SIP telephony, and certificates needed for authentication to a service provider that would otherwise be required to be implemented by the device locally. Low manufacturing costs potentially allow consumers to deploy videoconferencing on every television in the home. Finally, since DCAs are connected to TVs through an auxiliary or similar type of connector, the DCA and the videoconferencing experience is selected from the TV remote like any other channel.

Figure 3 shows a simplified representation of the functional blocks that compose the DCA. Thick black lines indicate the flow of audio and video media. Thin black lines indicate the flow of control information.

The A/V stream control module simultaneously manages the flow of received packets of audio and video, as well as transmitted packets. An external microphone and video camera captures audio and video. If the sensors are analog, analog to digital conversion (A/D) modules are be required to generate digitized samples. Digital to analog conversion (D/A) is additionally required to support analog output devices.

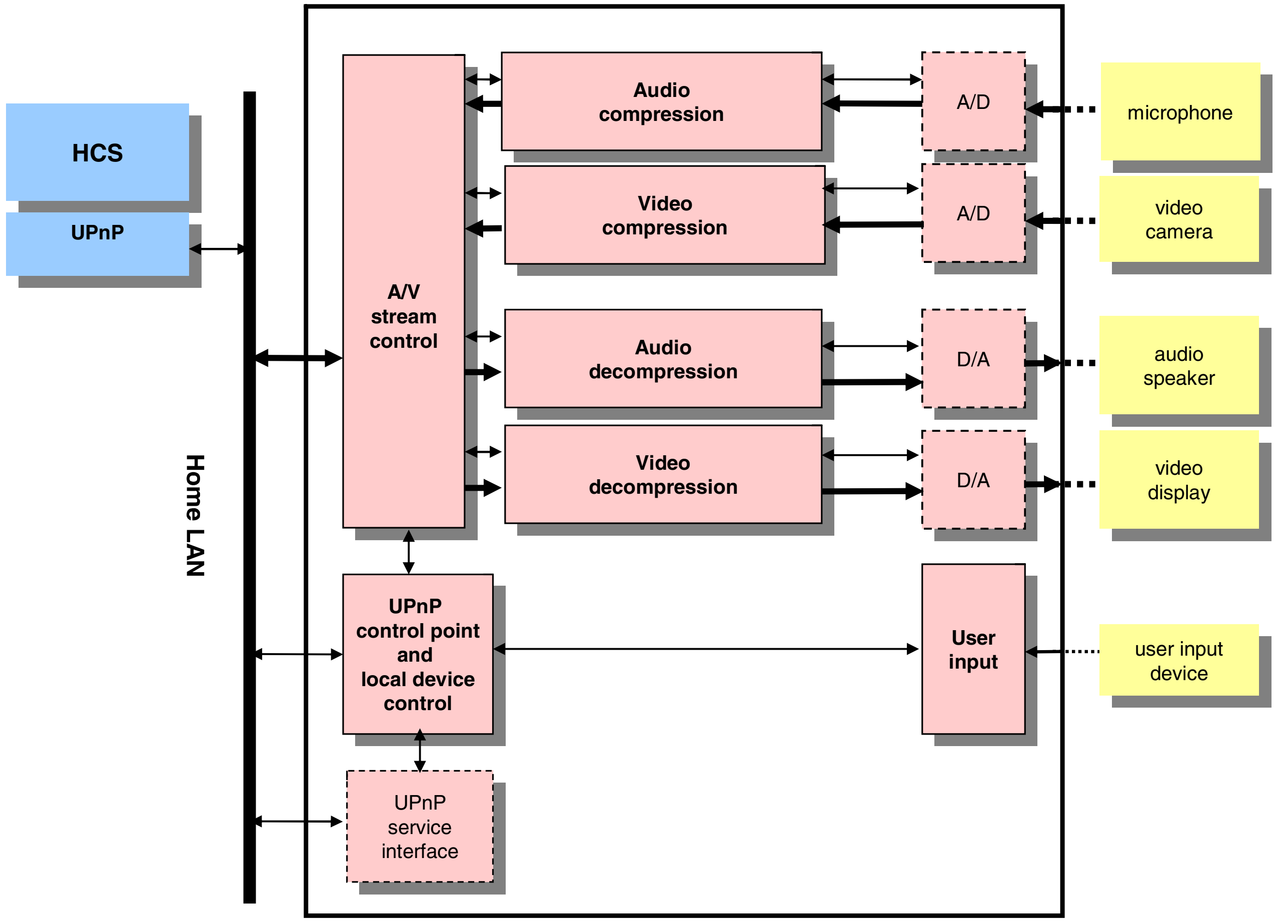

Figure 3: Functional blocks composing the DCA

The UPnP control point and local device control module allows the DCA to discover and utilize the HCS and manage commands received from the DCA user input device. Figure 4 shows a DCA connected to the home network. When an HCS is located, the DCA utilizes the browse contacts capability to identify contacts with videoconferencing capability. Contact metadata include sufficient information to establish media format and transport protocol compatibility between caller and callee in a manner transparent to the user.

A DCA implementation may optionally host its own UPnP service interface (dotted line module at lower left in Figure 3), enabling the DCA to be utilized by other UPnP devices connected to the LAN. The DCA UPnP services may allow the DCA to be notified of and receive inbound calls from other communication devices inside the home.

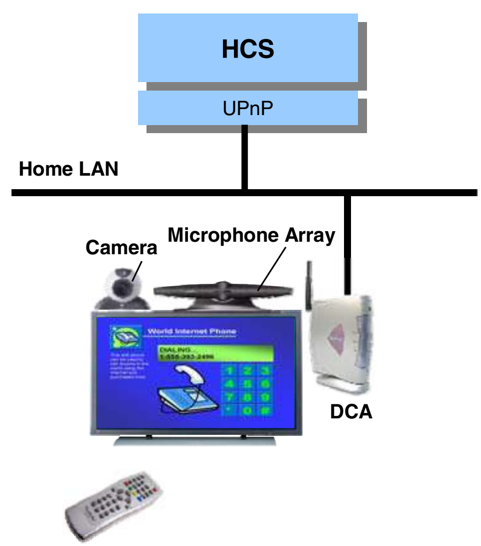

Figure 4: DCA connected to the home communications network

Meeting the Need for a Greater Sense of Closeness

High-definition videoconferencing and the option to persist video sessions over extended periods of time effectively maximizes the use of technology to achieve a greater sense of family closeness.

THE DLNA FRAMEWORK–ROLE MODEL FOR A DIGITAL HOME COMMUNICATIONS FRAMEWORK

Consumers are acquiring, viewing, and managing an increasing amount of digital media on devices in the CE, Mobile Device, and PC domains. Consumers want to conveniently enjoy that content across different devices and locations in their homes. Delivering this vision mandated the creation of a common set of industry design guidelines to ensure interoperability and additionally to allow companies to participate in a growing marketplace.

As a result, the DLNA Home Networked Device Interoperability Guidelines 1.0 were created. The Interoperability guidelines are based on an architecture that defines interoperable components for devices and software infrastructure. The DLNA architectural components are as follows:

- physical media
- network transports
- device discovery and control
- media management and control
- media formats
- media transport protocols

The essential technology ingredients that underlie the architectural framework for digital home entertainment in the Interoperability guidelines are generally the same technology ingredients required to establish a common architectural framework for digital communication devices: Ethernet, 802.11, IPv4, HTTP, and UPnP technology. A successful interoperable framework for communications would therefore follow the example of the DLNA by institutionalizing an industry-wide consensus on the following:

1. Essential technologies (in addition to the ones enumerated above) for home communications.
2. Common architectural components in communication applications, both present and future.
3. Common transport protocols.
4. Common audio and video formats.

CONCLUSION

The results of a recent CMR study of unmet consumer communication needs were summarized. Unmet communication needs discovered in the study were grouped into the categories of Continuity, Management, Protection, and Closeness. Analysis of the needs data suggests that the physical home is a good base for deploying technology solutions that combine personal communication devices in a manner that affords users a feeling of home while they are away from home. A user-transparent, home framework for connecting different communication devices to each other, both inside and outside the home addresses the problem of bridging disparate communication channels. Enabling full user personalization and control of wanted and unwanted contacts, personalized policies for call and message handling, etc. addresses the consumer desire for communication protection. Richer communications experiences such as videoconferencing or sharing pictures of the family while talking help address the need to achieve a greater sense of closeness.

A new framework for digital home IP communications possessing architectural components similar to those of the established DLNA framework is proposed. Two new device classes that exploit the framework are additionally proposed. A Home Communications Server is proposed to satisfy user needs for communications continuity, management, protection, and a greater sense of closeness.

The HCS will enable a new class of devices that depend on its standardized network APIs. Digital Communications Adaptors are devices in this category and address the unmet need for a greater sense of closeness by performing personal videoconferencing and media sharing during a call. The DCA depends on the HCS to respond and connect it to incoming calls. DCAs achieve low implementation cost by off-loading functionality to the HCS such as contact archives, SIP telephony, and certificates needed for authentication to a service provider. DCAs are connected to TVs through an auxiliary or similar type of connector, allowing the DCA and the videoconferencing experience to be selected like a TV channel.

ACKNOWLEDGMENTS

We acknowledge Antony Barton and the members of the Corporate Market Research team for providing the CMR data and for granting permission for its use in this paper.

REFERENCES

[1] *Home Networked Device Interoperability Guidelines v1.0, 2004*, www.dlna.org.*

[2] Olga Kharif, "VoIP Providers: Heeding the Call?" *Business Week Online**, November 28, 2005.

[3] "Signaling flows for the IP multimedia call control based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3," *3GPP TS 24.228*, www.3gpp.org/ftp/Specs/html-info/24228.htm*

[4] "IP Multimedia Subsystem (IMS); Stage 2," *3GPP TS 23.228*, www.3gpp.org/ftp/Specs/html-info/23228.htm.*

[5] "SIP: Session Initiation Protocol," *IETF RFC3261*, June 2002, www.ietf.org/rfc/rfc3261.txt.*

[6] "Session Initiation Protocol (SIP): Locating SIP Servers," *IETF RFC 3263*, June 2002, www.ietf.org/rfc/rfc3263.txt.*

[7] "The Session Initiation Protocol (SIP) UPDATE Method," *IETF RFC 3311*, Sept. 2002, www.ietf.org/rfc/rfc3311.txt.*

[8] "A Message Summary and Message Waiting Indication Event Package for the Session Initiation Protocol (SIP)," *IETF RFC 3842*, Aug. 2004, www.ietf.org/rfc/rfc3842.txt.*

[9] *UPnP Design by Example*, M. Jeronimo, J. Weast, Intel Press, 2003.

[10] "SDP: Session Description Protocol," *IETF RFC 2327*, April 1998, www.ietf.org/rfc/rfc2327.txt.*

[11] "HTTP Over TLS," *IETF RFC 2818*, May 2000, www.ietf.org/rfc/rfc2818.txt.*

AUTHORS' BIOGRAPHIES

Mark Walker is a senior architect in the Intel® Viiv™ Technology Platform group. Mark joined Intel in 1993 and his past duties have included research and development in the areas of video/audio/speech compression, text-to-speech, videoconferencing, and digital packaging. More recent activities include contributing architect and chair positions in ITU, ISO, UPnP, and CEA standards efforts focused on developing an interoperable framework for digital home entertainment. Mark is the author of eight patents in the areas of video, audio, and speech compression, IP-based communications, advanced Internet application networking, and home Ethernet. Mark holds B.S., M.S. and PhD degrees in Electrical Engineering, all from Arizona State University. His e-mail is mark.r.walker at intel.com.

Jim Edwards is a principal engineer and manager in Platform Architecture and Advanced Development in Hillsboro, Oregon, working on the architecture, standards, and tools/stacks for consumer PC platforms in the digital home. He oversees a team of developers and architects responsible for leading the development of key digital home standard specifications. Jim was previously the chief architect for Intel's Extended Wireless PC initiative. Prior efforts that Jim participated in at Intel Labs led to the development of protocol independent telephony and network APIs for the PC. He holds five patents in communications and networking. Jim holds B.S. and M.S. degrees in Electrical Engineering from Cornell University. His e-mail is jim.edwards at intel.com.

Ylian Saint-Hilaire is a senior software engineer and architect at Intel Research Labs in Hillsboro, Oregon. Ylian began his Intel career in network security, focusing on enterprise-wide deployment of IPsec and VPN. Ylian then became an architect in the Intel Digital Home Labs leading the development of Intel UPnP technologies which became the leading UPnP compliant stack. Ylian also worked with the UPnP forum on the definition of remote displays, and he worked with the Intel team on the development of digital media adaptors. He holds B.S. and M.S. degrees in Computer Science from the University of Quebec in Montreal (UQAM), and he holds four patents in communications and networking. His e-mail is ylian.saint-hilaire at intel.com.

Jeffery Lee is a strategic planning manager in the Digital Home Group, where is he responsible for long-range plans, market/business analysis, technology definition, strategies, and roadmaps for consumer platforms. Jeff has been with Intel since 1990 and is located in Folsom, California, where he has held a variety of roles in strategy, planning, product marketing, and technical marketing working on flash memory, cellular/handheld devices, graphics, and chipsets. Before coming to Intel, Jeff was with Oracle Corporation and Xerox. Jeff holds a B.S. degree in Computer Science from the University of California, Berkeley, and a Masters in Business from California State University, Sacramento. His e-mail is jeffery.lee at intel.com.

® ™ Intel and Intel Viiv are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

THIS PAGE INTENTIONALLY LEFT BLANK

Seamless Collaboration–Enabling Best-in-Class VoIP Experience on Intel® Centrino® Mobile Technology

Hani Elgebaly, Mobility Group, Intel Corporation
Farid Adrangi, Mobility Group, Intel Corporation
Rajeev Muralidhar, Mobility Group, Intel Corporation
Lakshmi Ramachandran, Mobility Group, Intel Corporation
Manish Nair, Mobility Group, Intel Corporation

Index words: Voice over Internet Protocol (VoIP), seamless collaboration, wireless LANs, Quality of Service (QoS)

ABSTRACT

With the introduction of Intel® Centrino® mobile technology, today's mobile notebook is an advanced communication platform with integrated wired and wireless technologies. One of the important communication technologies that has been making excellent progress in the industry is Voice over IP (VoIP). The Centrino notebook augments ingredient technologies to improve the VoIP experience for the end user, resulting in enhanced productivity.

Recent advances in VoIP applications, protocols, infrastructure, equipment and peripherals have resulted in IT departments of medium and large businesses becoming interested in deploying this technology. The goal is to have the business road-warrior communicate in a seamless way both inside and outside the enterprise. Also with the advancements in the Internet Protocol Private Branch Exchange (IP PBX) infrastructure, IT shops have been racing to replace their legacy, cumbersome Time Division Multiplexing (TDM) PBXs with easy-to-manage and more flexible IP PBXs. The impediments to such deployment have been the ability to deliver the expected Quality of Service (QoS) for voice calls, call privacy, and ease of use of VoIP applications.

In this paper, we present the platform ingredient technologies that Centrino offers this year (2006) in order to deliver a best-in-class VoIP experience over wireless and wired networks. This includes Intel's QoS solution for enhanced VoIP experience over Wireless Local Area Networks (WLAN), Intel's Array Microphone for better open audio, Intel's Integrated Performance Primitives (IPP) libraries for improved overall VoIP experience, and enabling the use of Bluetooth* headsets for hands-free VoIP.

INTRODUCTION

VoIP is being increasingly deployed by enterprise IT departments as a cost-effective replacement for PBX-based telephone networks. The single biggest advantage of deploying VoIP is that IT administrators have to maintain just one kind of network for both data and voice applications resulting in significant cost savings. Moreover, the quality of VoIP calls has improved significantly over the past few years as voice quality-related issues have been addressed in data networks as well as in VoIP devices. VoIP is also gaining popularity in the consumer space, in homes and Wi-Fi public hotspots like hotels, airports, and coffee shops.

The popularity of WLANs has tremendously increased among notebook users with ubiquitous Wi-Fi deployments in enterprise networks and popular places like airports, hotels, and coffee shops. Enterprises of all sizes are adopting WLANs as connectivity extensions of their corporate network beyond office spaces and cubicles, or in some cases as a replacement for the wired network.

As VoIP and WLAN deployments continue to gain momentum in the enterprise, VoIP usage over WLAN is also expected to gain popularity, especially among notebook users with softphone applications. Supporting adequate quality for VoIP over WLAN has some unique challenges arising from the characteristics of WLANs. Typically, the bandwidth available over a WLAN is an order of magnitude less than that available over a wired LAN. Moreover, a WLAN signal is susceptible to interference from adjacent WLANs as well as other radio devices. This causes increased delays and sometimes losses for packets transmitted over WLAN. Another big challenge is that of Quality of Service (QoS) for VoIP traffic. Intel's Seamless Collaboration architecture, described in this paper, addresses WLAN QoS challenges and introduces other platform technology ingredients to improve the VoIP experience on Intel Centrino mobile technology notebooks.

In this paper, we first provide an overview of Intel Centrino 2006 Seamless Collaboration scenarios. We then describe the network topology and landscape for VoIP and its deployment challenges in enterprise networks, and for VoIP QoS over wired and wireless networks. In the following sections, we present Intel's QoS architecture for VoIP over WLAN, and we describe the Intel Integrated Performance Primitives (IPPs) that are available for softphone optimization. We then provide an overview of Intel's Array Microphone architecture and Bluetooth-Wireless Coexistence solution for Bluetooth headset enabling.

VOIP SEAMLESS COLLABORATION USAGE SCENARIOS

Seamless Collaboration brings some new usage models to Intel Centrino mobile technology 2006 platforms with the focus on enhancements for integrated communication devices. These new usage models allow the notebook to become an effective communication device for voice anytime/anywhere. There are several hardware/firmware, software/drivers, and optimization components that support the Seamless Collaboration architecture.

There are three key usage scenarios that will be enabled by Seamless Collaboration: "Office on the Go," "Multi-party VoIP conferencing," and "Integrated communication experience."

Office on the Go

This enables users to use their Intel Centrino mobile technology notebooks as a central communication device for voice anytime/anywhere. The user can make and receive voice calls anywhere as long as he/she has IP network connectivity to his/her enterprise network, as if he/she were physically present at the office. Many users will be connecting to their respective IP networks via WLAN and will use wireless headsets/handsets using Bluetooth. This usage model also allows the user's calls to be automatically routed to the user's current location. For example, the user could have incoming calls to both his/her desk and cell phone be automatically forwarded to his/her notebook softphone, whenever connected to the enterprise network. The user would be able to configure which phones and calls are automatically forwarded to his/her current location. This way the user can effectively create a virtual office experience.

Figure 1: Office-on-the-Go scenarios

Multi-Party VoIP Conferencing

This feature enables users to use VoIP softphones for multi-party audio conferencing (including adhoc and bridge-based conferences) with optimized narrow-band and wide-band audio codecs and Intel Microphone Array open audio technology. This usage scenario allows a user to conference in more than two parties using the softphone application on the notebook, thus saving on expensive conference bridges. Dual core and other platform ingredient technologies enable hosting N multi-party conferencing sessions on Intel Centrino mobile technology, where N is greater than 8.

Integrated Communication Experience

One instance of this usage scenario provides users the facility to receive their voice messages as e-mail in their Outlook* inbox. This allows users to access their voice mail stored as e-mail in their inboxes and to listen to them in any order. Users can also choose to reply to the voice mail via a text message. Other instances of this usage scenario include integrated optimized softphones with productivity and e-business enterprise applications (Click-to-Dial scenarios). For example, if an employee in the company's Enterprise Resource Planning (ERP) system has a question or inquiry about a transaction, he/she can click to dial customer support and instantly talk with a support representative.

* Other names and brands may be claimed as the property of others.

Key VoIP Client Components

Softphone Application

A softphone is a software VoIP application that provides a phone-like user interface to a notebook user complete with a dialing pad on the screen that can be clicked with a mouse. The softphone application may interface with external peripheral devices such as Bluetooth or a USB microphone and/or headset. Some softphone applications may also allow a user to configure VoIP parameters such as voice codec, call characteristics, or data rate.

Most softphone applications require a call manager on the network that authenticates softphone users, maps phone extensions to IP addresses, and routes VoIP call signaling messages between softphone users. A call manager can run on a server for PC-to-PC calling. Alternatively, a call manager can be integrated with a hybrid PBX, in which case it can also provide a connection to the enterprise and external PSTNs.

Audio Peripherals

A softphone application uses notebook audio peripherals such as codecs, built-in speakers, and microphones to convert user speech into VoIP packets and vice versa. Alternatively, USB or Bluetooth peripherals, such as handsets, headsets, speakers, and microphones, can also be used as audio peripherals.

Audio Codecs

A software audio coder-decoder (codec) is needed to sample and encode audio from a microphone to bits that can be sent in a VoIP packet and to decode bits from incoming VoIP packets to audio signals that can be played through the speakers or headphones. A variety of codecs is available for different conditions such as available bandwidth, audio quality, protection from lost packets, etc. Most audio codecs developed in the last few years were either designed to provide quality comparable to PSTNs (e.g., G.711 [8]) or to allow the use of VoIP with low-bandwidth networks such as dial-up (e.g., G.729 [8]). These codecs typically sample human voice signals at 8 kHz to allow representation of audio frequencies up to 4 kHz. Codecs with sampling rates up to 8 kHz are known as narrowband codecs. More recently, wideband codecs have been developed with sampling rates as high as 16 kHz, which typically result in better perceived audio quality than is possible with narrowband codecs, because audio frequencies up to 7-8 kHz can be adequately represented. The use of such codecs is however limited to PC-to-PC VoIP calls, because a PSTN typically cannot transport or reproduce audio frequencies higher than 4 kHz.

Each codec can work with one or more frame rates, e.g., 20 ms, 30 ms, etc. A frame rate represents the time interval for which converted audio bits are encapsulated in an RTP packet to be sent to the other VoIP endpoint. Depending on the encoding algorithm, a codec produces either a fixed size or a variable size (in bits) audio sample. The data packet transmitted over a LAN or WLAN also consists of a User Datagram Protocol (UDP), an Internet Protocol (IP), a Logical Link Control (LLC), and Media Access Control (MAC) headers.

VOIP DEPLOYMENT IN ENTERPRISE NETWORKS

Traditionally, enterprises have maintained two separate networks: a voice network, based on PBX and PSTN and an IP network for data applications, such as email, Web, VPN, etc. VoIP technology allows the separate voice and data networks to be merged into a single network because voice can be treated as just another application running over the IP (data) network. Traditional PBX and phone extensions can still be supported using hybrid PBXs that can manage traditional (analog) phones as well as VoIP phones connected via Ethernet. Many enterprises have chosen to deploy a single IP network in their new office buildings where there is no need to support existing analog (PSTN) phone extensions. In addition, many enterprises have replaced their analog phone networks with VoIP [9].

Figure 2 shows an example enterprise network with desktop and notebook clients connected to LAN and WLAN segments. The WLAN segment is operated by an IEEE 802.11 Access Point (AP). These segments are connected to the corporate network via a wired backbone. Other components on the corporate network shown are enterprise servers, a hybrid PBX, legacy phones, and VoIP phones. In addition to these components, call managers are responsible for registration and authentication of VoIP clients and for routing VoIP signaling messages to the correct VoIP endpoints. Call managers also manage phone extensions mapped to VoIP phones (soft and hardphones). A media gateway, on the other hand, converts voice samples from VoIP packets on the VoIP network to a modulated analog signal used by the PSTN.

Beyond the AP, the enterprise network comprises a WLAN Distribution System (DS)–typically implemented as an Ethernet-based LAN, Layer 3 routers, switches, IP / Multi-Protocol Label Switching (MPLS)/Differentiated Services (DiffServ) domains, etc. Networks at multiple sites may be interconnected via VPN/MPLS tunnels/frame relay to give the appearance of a single corporate network. A VoIP call initiated from a WLAN client might terminate at an end host that is outside the Extended Basic Service

Set (BSS) serviced by the WLAN DS. In order to provide service differentiation across the entire end-to-end VoIP data flow path, the AP and each router/switch on the way must be configured to identify and prioritize VoIP flows. Details of QoS support required in various enterprise network components are given in a later section.

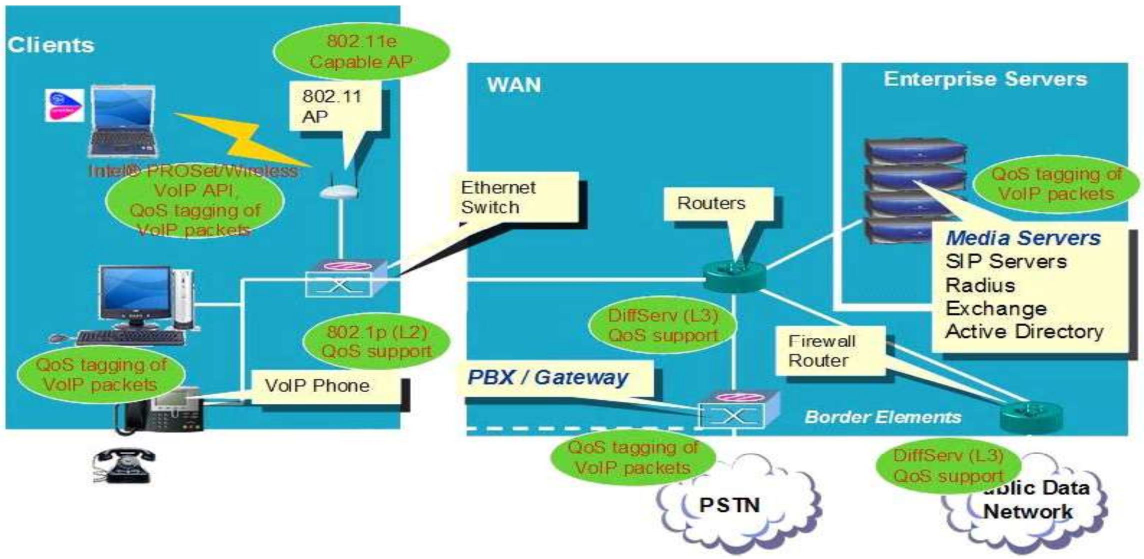

Figure 2: Enterprise VoIP deployment across wired and wireless networks

VoIP QoS over Wired Network

In order to provide service differentiation over the end-to-end VoIP data flow path, each router/switch on the way must be configured to identify and prioritize VoIP flows. This can be achieved as follows using Differentiated Services Code Point (DSCP), and it is also shown in Figure 2.

1. VoIP client endpoints must mark the VoIP packets that they send out with a pre-configured DSCP value.
2. In a large enterprise network consisting of multiple campuses/sites, when a VoIP packet arrives at a border router in the VoIP sender's network, it is identified as a VoIP packet from the DSCP in the IP header. The border router may use another DSCP value to ensure that this packet is given a prioritized treatment in the core enterprise network or VPN/Frame Relay tunnels that connect multiple sites.
3. When the VoIP packet reaches a border router in the receiver's network, it is identified as a VoIP packet based on the DSCP marking used in the core enterprise network. The border router may change the DSCP to a value that is reserved for VoIP flows in the receiver's network. The packet is then forwarded towards the receiver.

An IP PBX can use DSCP tags on its signaling packets so that these are forwarded with a higher priority over best-effort traffic in the enterprise LAN. If a SIP server is used for signaling, it can also send SIP signaling packets with the right QoS tags. A media gateway acts like another VoIP endpoint as far as the data network is concerned and hence it also tags the VoIP packets it sends with the appropriate DSCP values.

VOIP QOS OVER WLAN

The first published IEEE 802.11 WLAN standard [1] did not have any provision for supporting QoS over the wireless medium. This problem was subsequently addressed by the draft IEEE 802.11e standard [2]. In addition, an industry body known as Wi-Fi Alliance, formed for certification of IEEE 802.11 standards, published its own interoperability specification called Wireless Multimedia (WMM) [3], which is based on an earlier draft version of IEEE 802.11e. QoS support in IEEE 802.11e comes in two flavors:

1. **Prioritized QoS:** This allows classification of WLAN traffic into different categories based on their priorities. Higher priority traffic is given preferred access to the WLAN over lower priority traffic. This is achieved by a channel access function known as Enhanced Distributed Coordination Function (EDCF), or Enhanced Distributed Channel Access (EDCA). Prioritized QoS only provides a statistical guarantee. Voice and network control traffic are given the highest priority followed by video, best-effort data, and background data.
2. **Parameterized QoS:** Support for parameterized QoS is provided using a centralized Hybrid Coordination

Function (HCF) at the WLAN AP. HCF allows the wireless medium to be alternately used for contention-based access (using EDCF) and contention-free access (using HCF controlled channel access or HCCA). In the HCCA scheme, the AP grants opportunities for the WLAN Stations (STAs) to transmit by polling them based on their traffic requirements.

Admission control can be used with both the above schemes to limit the type and mix of calls on the AP. Parameterized QoS is optional and is not widely supported by AP and WLAN client vendors. This paper refers to only the EDCA scheme of IEEE 802.11e/WMM.

To achieve priority-based packet processing, a VoIP client needs to mark the VoIP packets that they send out with a pre-configured DSCP value. This will ensure that VoIP packets are processed with high priority according to the WMM/IEEE 802.11e specification implemented by the MAC layers of the WLAN client and the AP.

If the receiver is also a WLAN client associated with an AP, IP routing and Layer 2 switching will deliver the VoIP packet to the AP. The AP identifies the packet as a VoIP packet from its DSCP marking. Additional classification based on IEEE 802.11e Traffic Classification (TCLAS) may be performed by the AP in its WLAN MAC layer to classify the packet to a particular Traffic Stream (TS) for an associated WLAN client. The VoIP packet is then scheduled for delivery to the WLAN client by the AP's MAC layer.

Priority-based packet processing using EDCA mechanisms definitely helps real-time traffic like VoIP, but this is not adequate to provide hard guarantees. For example, an EDCA-based WLAN cannot support an infinite number of VoIP calls even though VoIP traffic can be treated with higher priority. Based on channel conditions, existing load on the AP, etc. there is a limit on how many calls can be supported. Once this limit is reached, MAC collisions, retransmissions, etc. cause delays to all the VoIP calls, thereby degrading the call quality. Hence, call admission control is needed. Once the AP enforces a limit on the number of VoIP calls that can be supported as high-priority calls, all subsequent requests must be rejected in order to maintain the delay/bandwidth guarantees expected for the already admitted calls. More details on how prioritization and admission control work in WLANs are available in IEEE 802.11e/WMM.

Our simulations show that both priority packet processing and call admission control are essential for providing call capacity and bandwidth guarantees to VoIP calls. Figure 3 shows the improvement in VoIP capacity with EDCF, over plain IEEE 802.11 WLANs, for a single AP. The effect of admission control on the quality of the VoIP calls can be seen in Figure 4, where the quality of admitted calls does not go down when traffic exceeds the capacity limit. Calls that are denied admission are treated as best effort, so that admitted calls continue to get the reserved QoS.

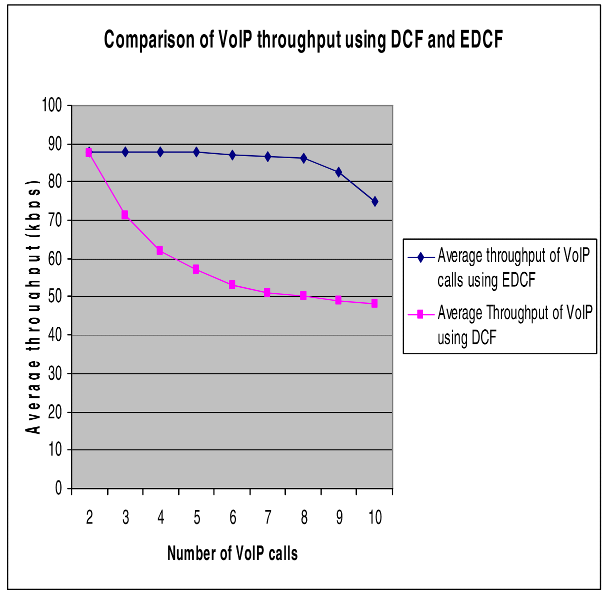

Figure 3: Comparison of performance of VoIP call in DCF vs. EDCF

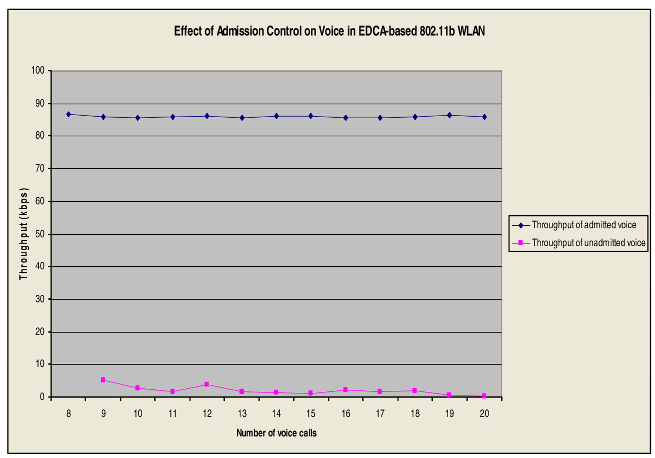

Figure 4: Effect of call admission control on the performance of VoIP calls

VOIP OVER WLAN CLIENT ARCHITECTURE

Figure 5 shows the client side of Intel VoIP over WLAN architecture, which consists of the following key components:

1. Softphone application.

2. Intel WLAN Softphone Application Programming Interface (API): Softphone applications use this interface to convey VoIP profile information to the WLAN QoS Manager. A *VoIP profile* is a collection of application-level parameters that defines the QoS requirements of the application. This API provides profile and flow management functionality. In addition, this can also be used to collect statistics for a VoIP flow such as packet loss, delay, and jitter, which can be used by the application for troubleshooting or for renegotiating the call.

3. WLAN QoS Manager: The WLAN QoS Manager is the interface between the softphone application and the WLAN driver. This is represented by the "Winsock Layer/GQoS API" block in Figure 5. It provides translation of a VoIP profile into appropriate parameters as understood by the WLAN MAC layer, i.e., traffic specification (TSPEC) and traffic classifier (TCLAS) information as defined in IEEE 802.11e/WMM.

4. WLAN MAC Layer: The Client WLAN MAC layer is split across the WLAN driver and the WLAN NIC. It contains both standards-based implementations of IEEE 802.11/.11e as well as any optimizations that Intel VoIP over WLAN architecture provides for improving the VoIP experience.

5. Intel Management Application: Intel management application can be used by IT managers to configure new VoIP profiles and to delete/modify existing VoIP profiles.

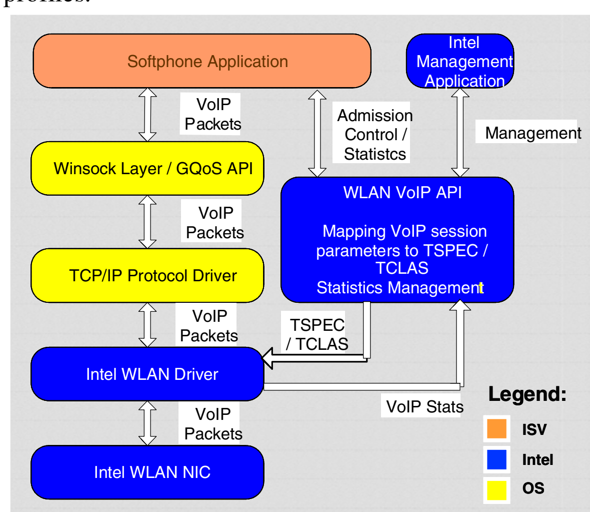

Figure 5: VoIP over WLAN client architecture

Note that VoIP and signaling packets generated by the softphone application (and its peer application/call manager) traverse the path shown through the Winsock [7] layer, TCP/IP stack, and the WLAN subsystem. The control path shown on the right side of the diagram is used to set up a VoIP flow over the WLAN network between the client and the AP.

Figure 6 shows an example sequence of VoIP flow creation in the Intel VoIP over WLAN architecture. When the softphone (VoIP) application is ready to initiate or accept a VoIP call, it first queries the list of VoIP profiles supported by the Intel WLAN VoIP API. From this list, the application picks a VoIP profile, depending on the codec and frame rate it has selected for the call, and it issues a *Create Flow* request to the VoIP API with the selected profile and a classifier, that describes the VoIP traffic. The VoIP API and QoS Manager map the VoIP profile to a WLAN Traffic Specification (TSPEC) and the classifier to WLAN TCLAS and request the WLAN driver to add a traffic stream over the air using these parameters. The WLAN driver sends the request to the AP according to the protocol specified in WMM/IEEE 802.11e. If the request for Create-Flow is accepted, the AP allocates medium time to the VoIP call and returns a successful response to the client. Otherwise, the AP simply returns a failure response to the client. The WLAN driver then processes the response and returns a success or a failure to the QoS Manager for the Created-Flow request. Lastly, the QoS Manager makes a callback into the softphone application via the VoIP API to inform it of the result of the call creation.

In case of a failure response, the softphone application may generate a busy signal indicating to the user that resources for the call are not available. The user may try the call again at a later time.

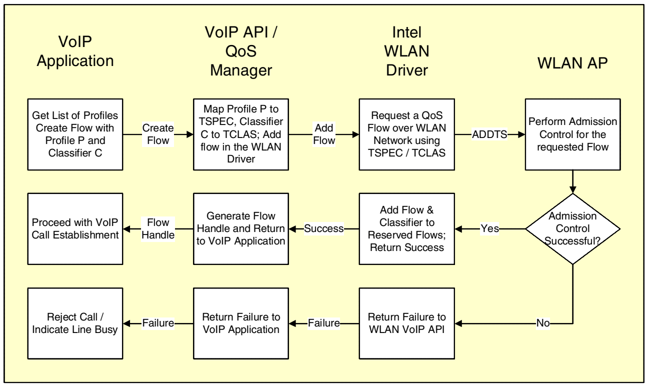

Figure 6: Example of VoIP flow creation over WLAN

VPN Implications

Most VPN technologies rely on "tunneling" the original data traffic from one VPN endpoint to another using some form of encapsulation. Since the original data packet is encrypted before it is transported over a VPN tunnel, header attributes from the original packet such as UDP/TCP port numbers, IP addresses etc. are not visible to a WLAN driver when it examines a VPN packet. This limits the ability of WLAN drivers to correctly identify and prioritize a flow. As an example, if an application creates a VoIP flow reservation in the WLAN driver and provides a classifier containing source/destination IP addresses, port numbers and protocol, the driver will be able to create the flow but will fail to identify packets that belong to this flow. Thus the VoIP packets will not be assigned to the correct flow reservation and will not get the priority over WLAN that they deserve. To avoid this, when a VPN connection is in use, the application should use only a DSCP value as a classifier instead of using a full transport layer classifier. When the application detects the presence of a VPN, it can use a classifier consisting only of tunnel source and destination IP addresses and/or DSCP value in the IP header instead of using a full transport layer classifier.

Today's VPN products give more importance to security than performance. Both L2TP and PPTP (two popular VPN tunneling protocols) bring in performance issues due to the processing and overhead involved in encrypting and encapsulating the packets. In PPTP, the packet is encapsulated inside a Generic Routing Encapsulation (GRE) packet, which is then encapsulated inside an IP packet before being sent across the tunnel. In L2TP, packets are encapsulated 4 to 6 times depending on the IPSec policy used. It also provides additional levels of security through the use of DES/3DES encryption which impacts performance.

To avoid performance and classification problems in the WLAN driver, it is highly recommended that enterprises use the IEEE 802.11i/WPA2 method instead of VPNs to encrypt WLAN traffic within the enterprise.

INTEL INTEGRATED PERFORMANCE PRIMITIVES

Intel IPPs are a set of optimized cross-platform software functions that boost media processing performance by taking advantage of processor microarchitecture, algorithmic techniques, and instruction sets such as Streaming SIMD Extensions (SSE3). Figure 7 depicts IPP components that could significantly enhance PC applications such as Audio, Video, Image, Graphics, Speech, Math, Signal, and Cryptography and save developers from costly and time-consuming hand-coding and optimization.

Table 1 below depicts a list of speech coding samples built with Intel IPPs as the building blocks that are bit-exact with the standard.

Table 1: List of codecs supported by Intel IPPs

Speech Coding Samples	Windows*	Linux*
G.722.1	✓	✓
GSM/WMR WB / G.722.2	✓	✓
G.723.1	✓	✓
G.726	✓	✓
G.728	✓	✓
G.729	✓	✓
GSM-AMR	✓	✓
GSM-FR	✓	✓

For more information on the Intel IPP Application Programming Interface (API), please visit http://www.intel.com/cd/ids/developer/asmo-na/eng/dc/mobile/242763.htm.

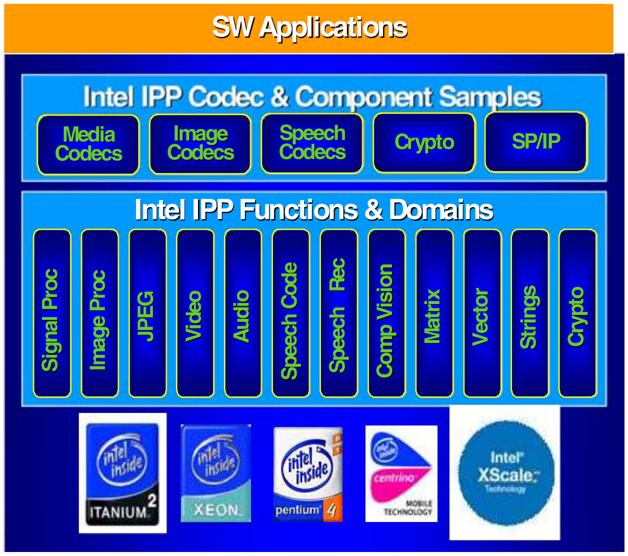

Figure 7: Software development with IPP

Figure 8 shows the performance gains of the highly optimized Intel IPP functions over compiled C code. See [Intel IPP Performance] for more information regarding Intel IPP performance. For more information about Intel IPP 5.0 including code samples, free evaluation copies, or to purchase copies visit [Intel® Integrated Performance Primitives 5.0].

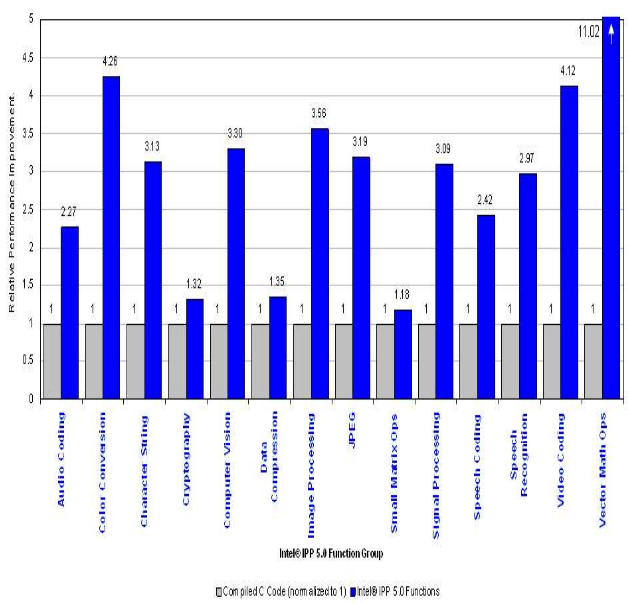

Figure 8: Intel IPP 5.0 Performance Primitives

HIGH-DEFINITION AUDIO AND ARRAY MICROPHONE FOR OPEN AUDIO

While other audio implementations have limited support for simple array microphones, Intel HD Audio supports larger array microphones. By increasing the size of the array microphone, users get incredibly clean input through better noise cancellation and beam forming. This produces a higher-quality input to voice recognition, VoIP, and other voice-driven activities.

An array microphone provides a speakerphone-like "open audio" usage solution for softphone VoIP calls. It does not require the user to wear headphones or a headset microphone; rather the user can speak directly into the device and listen to the platform speaker at the same time. The array microphone provides higher-quality audio and an improved user experience.

An array microphone is a set of multiple microphone elements integrated on the mobile platform to provide better microphone input quality and features, such as Acoustic Echo Cancellation (AEC) and repetitive noise attenuation and filtering or Noise Canceling .

Figure 9 depicts Intel Array Microphone architecture with four microphone elements.

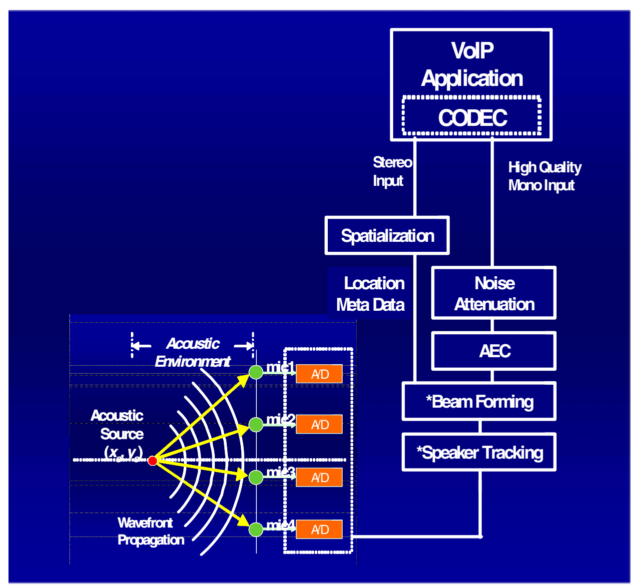

Figure 9: Intel Array Microphone architecture

Figure 9 depicts four key processes associated with Array Microphone processing. Speaker Tracking allows tracking of speaker source and direction to be able to replicate directionality at the rear endpoint if needed. Beam Forming implements speaker localization and sound attenuation. AEC ensures isolation between the speaker and the microphone for elimination of annoying echoes during the conversation. Finally, Repetitive Noise

Attenuation and Filtering ensures effective noise cancellation for clearer audio speech.

For more information about Intel High Definition Audio design and implementation please refer to http://www.intel.com/design/chipsets/hdaudio.htm.

Bluetooth Headset for Hands-Free VoIP

Bluetooth headsets are commonly used as an audio device with softphones. Bluetooth headsets enable users to have a hands-free VoIP conversation and to manage their connection, for example, accept/terminate a call, adjust volume, etc. A few softphone vendors have chosen to integrate Bluetooth headsets to enable ease of use with wireless headsets. The softphone and Bluetooth stack integration architecture on Microsoft Windows* operating system is depicted in Figure 10.

The Bluetooth stack is composed of the API interface; middleware protocols, such as the Service Discovery Protocol (SDP); Logical Link Control and Adaptation Protocol (L2CAP); RFComm (for serial communication); and Bluetooth profiles such as the headset profile, which is used for wireless headsets.

The softphones need to be integrated with the Bluetooth API to achieve service discovery and to interact with Bluetooth headset devices, i.e., be able to send and receive commands to and from Bluetooth headsets. Intel has worked closely with the Bluetooth stack and softphone vendors to define a common profile for VoIP control commands for headsets/handsets. Going forward, Intel will work with the Bluetooth stack vendors to define a common API for both headsets and handsets to enable softphones to interact with Bluetooth stacks from different vendors in the same way. This will significantly simplify the softphone integration efforts to support Bluetooth headsets for VoIP.

The coexistence of Bluetooth and WLAN radios on a notebook could impact the VoIP quality due to the interference between the two radios. Intel Centrino 2006 platforms implement the Wireless Co-existence Solution (WCS) to solve the problem of Bluetooth-WLAN coexistence by sending WLAN channels to Bluetooth. The WCS describes the interface between two radios running simultaneously in the same laptop to transfer information on the channel occupied by WLAN to a Bluetooth NIC. As a result, Bluetooth will skip the

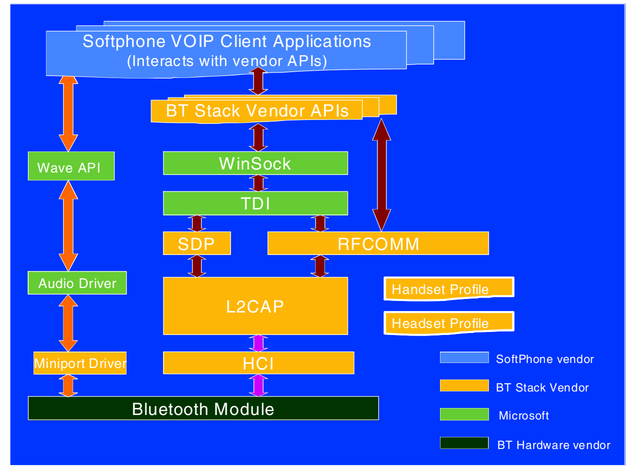

Figure 10: Softphone and Bluetooth stack integration

WLAN channel for non-critical events, and WLAN performance will be recovered in such cases. This solution significantly increases the throughput of WLAN when Bluetooth is presented. Additionally, the WCS specification provides better protection of critical Bluetooth communications (voice for example). This is accomplished by the following method:

- When Bluetooth expects to receive or transmit a high-priority packet on a Bluetooth channel that is located inside the WLAN channel, Bluetooth raises a BT_Priority signal.
- WLAN radio will then defer or kill its transmission if a BT_Priority signal is raised to avoid collision with Bluetooth packets.

For more information about Intel WCS solutions, please refer to [6].

CONCLUSIONS AND FUTURE WORK

In this paper, we presented Intel's Seamless Collaboration Architecture for VoIP on WLANs. We described our QoS architecture for VoIP that enables softphone applications to take advantage of QoS features on the Intel Centrino mobile technology notebook over WLANs. We also described other hardware/software solution ingredients, such as array microphones, IPPs, and the Bluetooth wireless coexistence solution, all of which make Intel Centrino mobile technology notebooks provide a best-in-class experience for VoIP over WLANs.

In addition to VoIP, collaboration usages that involve simultaneous audio, video, and data conferencing are becoming common in enterprises. Our future work is directed at extending this architecture to support multimedia collaboration applications. This includes hardware and software solutions for video peripherals like embedded cameras and remote handheld devices. Other usage models like multiparty ad hoc conferencing and IPP

* Other names and brands may be claimed as the property of others.

extensions for video conferencing are also part of this effort.

ACKNOWLEDGMENTS

We thank Ajay Bakre for his invaluable contributions to this work.

We also thank Duncan Glendinning, Selim Aissi, Ajay Gupta, Venkat Gokulrangan, Vijay Rao, Charlie Lasswell, and Sanjay Bakshi for the quality time they spent on reviewing this document and providing invaluable feedback.

REFERENCES

[1] *IEEE. Part 11*, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," ANSI/IEEE Std 802.11, 1999 Edition, (ISO/IEC 8802-11:1999(E)).

[2] *IEEE. Part 11*, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS), ANSI/IEEE Std 802.11e," Draft 13.0, January 2005.

[3] *Wi-Fi Alliance*, "WMM including APSD Specification," *WMM-Specification-draft-050720c.doc*.

[4] Rajeev Muralidhar, Ajay Bakre, Hani Elgebaly, "On the Performance of VoIP in EDCA-based IEEE 802.11 WLANs," in *Proceedings of World Wireless Congress, 2005*.

[5] http://www.skype.com.*

[6] Gordon Chinn et al., "Next-Generation PC Platform Built on Intel® Centrino® Mobile Technology New Usage Models," *Intel Technology Journal*, Q1'05.

[7] http://www.microsoft.com/msj/0599/LayeredService/LayeredService.aspx*

[8] http://www.itu.int/rec/recommendation.asp?type=products&parent=T-REC-g*

[9] Bernstein Research Call, Market Survey, "U.S. Telecom: Corporate Shifts to IP Driving Losing Battle for Carriers Within IT Budgets but Price Decays Moderating." May 2005.

AUTHORS' BIOGRAPHIES

Hani Elgebaly is the enterprise software architecture Manager in Intel's Mobile Platform Group. Hani's team is responsible for wireless and IP networking software technologies for VoIP, manageability, security, and mobile connectivity. Hani has been with Intel for ten years where he held multiple engineering management and lead architect positions at Intel labs (Corporate Technology Group) and with videoconferencing and IP telephony product groups. He has over 14 pending US patents and 15 journal and conference publications. He is also a senior IEEE member. He received his Ph.D. degree in Computer Science from the University of Victoria, British Columbia, Canada, his M.S. degree in Computational Science from the University of Saskatchewan, Canada, and his B.Sc. degree in Electrical and Computer Engineering with high honors from Cairo University, Egypt. His e-mail is hani.elgebaly at intel.com.

Farid Adrangi is an architecture manager in Intel's Mobile Platforms Architecture and Development Group where he leads architecture and development of Seamless Collaboration, one of the pillars in digital office initiatives. Farid's team is responsible for platform architecture and ingredients to deliver best-in-class audio/video/data conferencing experiences over wireless and wired networks. Prior to this, he worked in Intel's Research Communication Technology Lab as a lead architect on mobility and WLAN/WiMAX architecture and standardization efforts in 3GPP, IETF, and GSMA. Farid has two issued and over 15 pending US patents and 3 IETF RFC publications, and he has earned a GSMA Outstanding Award for his contribution to the Interworking WLAN Roaming architecture with 3GPP. He received a B.S. degree in Computer Science with high honors from Portland State University. He has been with Intel since 1994. His e-mail address is farid.adrangi at intel.com.

Rajeev Muralidhar is a senior software engineer in Intel's Mobile Platforms Architecture and Development Group where he works on architecture, research, and development of next-generation wireless software for Intel Centrino mobile technology platforms. Prior to this, he worked in Intel Research's Network Architecture Lab on architectures for network processor-based platforms for the Internet, including standardization efforts at the Network Processing Forum. Rajeev has over 6 pending US patents and over 10 journal and conference publications. He has an M.S. degree from Rutgers University, and a B.E. degree from NIT, India; both in Computer Engineering. He has been with Intel since 2000. His e-mail address is Rajeev.D.Muralidhar at intel.com.

Lakshmi Ramachandran is a staff architect in Intel's Mobile Platform Architecture and Development Group where she leads architecture, research, and development of next-generation wireless software for notebooks based on Intel Centrino mobile technology. In her 5 years at Intel, her areas of focus have included network management for SS7 stacks and network processors. Lakshmi has several research publications and over 5 pending US patents. She is also a co-author of a handbook on Mobile Wireless Networking, published by John Wiley and Sons. Her current focus areas are multimedia QoS for WLANs and new usage models for mobile platforms. She holds a Masters degree in Computer Science from the

Indian Institute of Science. Her e-mail address is Lakshmi.Ramachandran at intel.com.

Manish Nair is a senior software engineer in Intel's Mobile Platforms Architecture and Development Group where he works on architecture, research, and development of next-generation wireless software for Intel Centrino mobile technology platforms. Manish received his B.Tech degree from the College of Engineering, Trivandrum, India in Computer Science and Engineering. He is currently pursuing research in IP Security at the Indian Institute of Science, Bangalore. His e-mail is manish.nair at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

For further information visit:

developer.intel.com/technology/itj/index.htm