

SIP 기반의 홈서버용 영상전화 시스템 설계

안성호, 이경희, 이은령, 김지용, 김두현
한국전자통신연구원 컴퓨터.소프트웨어기술연구소
모바일협동작업연구팀
e-mail : ahnsh@etri.re.kr

A Design of Video Phone System based on SIP for HomeServer

Sung-Ho Ahn, Kyung-Hee Lee, Eun-Ryoung Lee, Ji-Yong Kim, Doo-Hyun Kim
Mobile Collaboration Research Team
CSTL/ETRI

요 약

VoIP(Voice over IP) 기술은 다양한 인터넷 응용 서비스 보급의 대중화에 기여한 주요 통신기술의 하나이다. 이에 따라 인터넷 이용자가 급격히 증가하고 인터넷 전화의 수요가 증가하게 되었다. 특히, 음성 뿐만 아니라 영상에 대한 기술이 접목되어 V2oIP(Voice and Video over IP) 라 일컬어 지는 기술이 보급되면서, 인터넷 영상전화에 대한 대중화가 이루어 지고 있다. 한편, 다양한 인터넷 응용 서비스 보급의 대중화에 따른 가정 내에서의 네트워크 환경 또한 부각되고 있어 정보가전분야에 큰 변화가 일고 있다. 이에 홈서버 중심의 홈네트워크환경이 구축된다. 따라서, 기존의 pc 를 단말기로 한 인터넷 환경 및 제반 응용 서비스가 그대로 홈서버 중심의 홈네트워크환경으로 옮겨져야 할 필요가 있다. 본 논문에서의 전체 시스템은 임베디드 리눅스 기반인 Qplus 운영체제를 기반으로 하는 홈서버상에 HoCoS(Home Collaboration Service) 시스템이 탑재되며, 이 시스템은 크게 영상전화 시스템과 공동작업 시스템으로 구성된다. 본 논문에서는 상기 시스템 중 SIP(Session Initiation Protocol)기반의 홈서버용 영상전화 시스템에 대한 설계 및 구현에 대해 기술하고자 한다.

1. 서론

인터넷 사용자 수의 급격한 증가와 다양한 인터넷 응용 서비스 보급의 대중화에 따라 인터넷에서의 VoIP(Voice over IP) 기술이 부각되어 널리 활용되고 있다. 특히, 음성 뿐만 아니라 영상에 대한 기술이 접목되어 V2oIP(Voice and Video over IP) 라 일컬어 지는 기술이 보급되고 있다. 한편, 다양한 인터넷 응용 서비스 보급의 대중화에 따른 가정 내에서의 네트워크 환경 또한 부각되고 있어 정보가전분야에 큰 변화가 일고 있다. 이에 VoIP 분야의 핵심 프로토콜 기술로는 H.323 프로토콜과 SIP(Session Initiation Protocol)가 있다. 현재 근거리통신망(LAN) 환경에서 멀티미디어 통신을 지원하는 H.323 프로토콜을 기반으로 하는 인터넷 전화, 영상회의 등에 적용해 왔으나 이 방식이 새로운 기술에 대한 확장성 및 호환성이 떨어지고 기술개발도 복잡하다는 단점이 있다. 이에 비해 최근에는 SIP 라고 불리는 프로토콜이 널리 대두되고 있고 많은 연

구가 활발히 이루어지고 있다[1]. 이 프로토콜 기술은 사용자 간에 멀티미디어 세션을 생성, 수정, 해제하는 응용 계층의 시그널링 프로토콜로서, H.323 에 비해 단순한 구조를 가지고 있고, 망 유지보수 및 관리가 편리하며, 다른 시스템과 확장성 및 유연성이 뛰어나 H.323 의 단점을 극복하고 새로운 대체 프로토콜로 부각되고 있다. 또한, 새로운 부가 서비스를 추가하기가 쉽다는 장점이 있어서 VoIP 서비스 분야의 확장이 예상된다. 한편, 인터넷 보급의 대중화에 따라 가정내의 네트워크 환경인 홈 네트워크가 주목 받고 있다. 이에 주요 분야로 정보가전(Information Appliance)이 핵심이며, 이는 홈서버 기반의 네트워크환경이 구축된다. 본 연구원에서는 임베디드 리눅스인 Qplus 운영체제를 기반으로 다양한 응용 서비스를 제공하고 있다. 가정내의 각종 기기에 대한 제어, 방법 및 감시, DTV 를 통한 여러 가지 통신 서비스 등이 있다. 이 중 본 논문에서는 SIP 기반의 홈서버용 영상전화 시스템에 대한 설계 및 구현에 대해 기술하고 있다.

본 논문의 구성은 다음과 같다. 2 장에서는 SIP 에 대한 특징 및 프로토콜 스택 구성도에 대해 설명하고, 3 장에서는 SIP 기반의 홈서버용 영상전화 시스템의 설계 및 시나리오에 대해 기술하며, 마지막으로 4 장에서는 결론으로 구성되어 있다.

2. SIP(Session Initiation Protocol)

2.1. SIP 개요 및 특징

SIP(Session Initiation Protocol)은 UAS(User Agent Server) 및 UAC(User Agent Client)간 또는 IP 폰과 같은 단말기 사용자들 간에 기존의 VoIP 서비스뿐만 아니라 다양한 서비스의 call 설정 프로토콜이다. 즉, SIP 은 P2P 시그널링 프로토콜이며, E-mail 과 유사한 주소 체계 형태의 동일 식별자(same identifier)를 이용하고, 언제, 어디서나 음성 통화 서비스를 비롯한 E-mail, IM(Instance Messaging) 서비스 및 Presence 서비스 등을 제공한다. 또한 SIP 을 이용하여 세션을 설정할 때 세션 파라메타를 협상함으로써 사용자의 단말기 성능에 따라 서비스가 다양하게 지원된다. 한편, SIP 은 텍스트 기반 프로토콜로서 HTTP 와 유사하며, 확장성이 우수하며, 사용자의 이동성(Mobility)을 용이하게 제공한다. UDP 및 TCP 에 상관없이 동작하므로 하부계층의 전송 프로토콜에 독립적이고, SDP(Session Description Protocol)를 세션의 정보를 협상하기 위한 메시지 바디로 사용한다[2].

2.2. SIP 프로토콜 스택 구성도

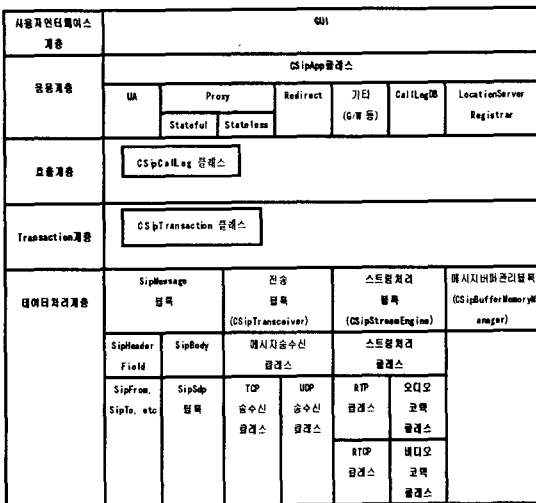


그림 1. SIP 프로토콜 스택 구성도

SIP 프로토콜 스택 구성도는 위의 그림 1 과 같다. 즉, 먼저 계층적으로 분류하였고, 각 계층에는 해당 블록으로 구성된다. 사용자 인터페이스 계층은 각종 응용 서비스에 대한 GUI 및 사용자로부터 어떠한 명

령을 받아들이는 최상위 계층이다. 응용 계층은 UA, Proxy Server, Redirect Server, Registrar 등 SIP 기반의 기본이 되는 응용 모듈이 위치하는 계층이다. 이 계층은 호출 계층, Transaction 계층, 메시지 계층을 관리하는 기능과 상대방과 메시지 송수신, 사용자 인터페이스간의 통신을 담당하게 된다. 호출(Call-Leg) 계층은 상대방과 호출을 시작하고 종료할 때까지의 상태를 관리하는 기능을 수행한다. 트랜잭션(Transaction) 계층은 transaction 의 상태를 관리하는 기능을 수행한다. 호출이 이루어지는 경우에는 호출 계층의 제어를 받지만 만약 호출이 아닌 transaction 의 경우에는 응용 계층의 제어를 받는다. 한편, 데이터 처리 계층은 SIP 메시지를 생성, 해석하는 기능, 메시지를 송수신하는 기능, 오디오/비디오 코덱 기능, RTP, RTCP 기능을 수행하는 계층이다.

3. SIP 기반의 영상전화 시스템 설계

3.1. SIP 기반의 영상전화 서비스

SIP 기반의 홈서버용 영상전화 서비스를 위한 기본 환경 구성도는 아래 그림 2 와 같다. 이는 2 차용 UA 와 유사한 구성을 한다. 영상전화 시스템은 사용자에게 기존의 전화 사용 환경을 유지하면서 인터넷 기반 영상전화 시스템으로 확장하도록 하는 것이다. 즉 원격지와 통신은 인터넷을 이용하지만 맥내에서 사용하는 단말은 기존의 가정용 유무선 전화기를 그대로 이용하도록 한다. 구현 형태는 홈서버에 영상전화 시스템이 탑재되고 일반전화기와 유무선 카메라 및 DTV 가 홈서버에 연결되어 있다.

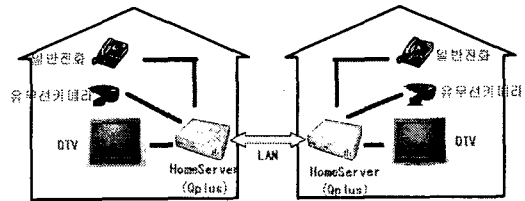


그림 2. 영상전화 서비스를 위한 환경 구성도

3.2. SIP 기반의 영상전화 시스템 구성

SIP 기반의 영상전화 시스템의 구성은 사용자 맥내의 홈서버를 중심으로 구성되는데, T.120 기반의 공동작업 시스템과 더불어 동작되기도 하고, 단독으로 동작하기도 한다. 참고로, 본 논문에서는 이를 통합 시스템으로 하여 HoCoS(Home Collaboration Service) 시스템이라 한다.

SIP 기반의 홈서버용 영상전화 시스템은 기본적으로 임베디드 리눅스인 Qplus 를 운영체제로 하는 홈서버에 탑재되며, 그 구성은 상대방을 호출하거나 상대방의 호출을 수신하는 부분과 일반전화기와 접속하는 부분, 유무선 카메라와의 인터페이스 및 Ethernet 인터

페이스, 그리고, 영상통신을 위한 H.261 소프트웨어 인코더 및 디코더, 음성통신을 위한 G.711 소프트웨어 인코더 및 디코더를 기본으로 한다.

그럼 1 을 참고로 SIP 기반의 홈서버용 영상전화 시스템의 구성을 살펴보면 다음과 같다.

- 사용자 인터페이스 및 GUI

이 계층은 사용자로부터 호출 명령을 받아 메시지를 생성하고 전달하는 기능, 상대방으로부터 호출 메시지를 수신한 경우 호출이 왔음을 알리는 기능 등 영상전화 서비스에 관련된 사용자와의 인터페이스를 관장하는 계층이며, 공동작업 서비스의 응용프로그램을 invoke 시키는 계층이다. 영상전화의 송수신 버튼 및 비디오 디스플레이 창이 있으며, 공동작업 응용프로그램을 invoke 시키는 버튼이 있다. 그리고, 상기 기능은 일반전화기의 * 및 #버튼을 이용하여 동작하도록 하였다.

- 영상전화부분

이 계층은 영상전화 시스템이 1 대 1 의 2 자간 UA 응용과 유사하므로 UA 의 기본적인 기능을 갖고 있다. 즉, 호출 계층, Transaction 계층, 메시지 계층을 관리하는 기능과 상대방과 메시지 송수신, 사용자 인터페이스를 담당하여 수행하는 계층이다. 먼저, 영상전화에서는 상대방과의 세션 연결 후 통신상태에서 영상전화 서비스가 이루어진다. 이러한 서비스를 지원하기 위해 영상관련 인코더 및 디코더 및 영상신호의 입출력 기능을 담당하는 영상모듈이 있고, 음성관련 인코더 및 디코더 및 음성신호의 입출력 기능을 담당하는 음성모듈이 있으며, 공동작업모듈간의 인터페이스 기능을 담당하는 SIP I/F 모듈 등으로 구성된다.

- 영상 모듈

영상모듈에는 영상신호의 입출력 및 압축 및 복원의 기능을 하는 클래스들이 있다. 영상신호의 입력 및 출력은 홈서버의 유무선 카메라의 API 및 드라이버 스펙에 따라 입력되며, 홈서버 별도의 모니터 혹은 DTV 를 통해 디스플레이 된다. 한편, 영상신호의 압축 및 복원은 ITU-T H.261 국제표준을 따른다. 화면크기는 현재 QCIF(176*144) 및 CIF(352*288) 이다. 일반전화기의 *버튼을 누르면 실행된다[3].

- 음성 모듈

음성모듈에는 음성신호의 입출력 및 압축 및 복원의 기능을 하는 클래스들이 있다. 음성신호의 처리는 ITU-T G.711 국제표준을 따르며, 홈서버에 연결된 일반전화기를 단말로 사용한다. 상대방의 ip 를 누르면 통화를 위해 세션이 열리게 된다.

- SIP I/F 모듈

SIP I/F 는 영상전화 모듈과 공동작업 모듈간의 통

신을 담당하게 된다. 그 내용은 최초 연결이 일어날 때와 최종 연결이 끊어질 때의 동기화에 대한 것이다. 실제로 영상전화 모듈은 구현 언어가 C (C++)이고, 공동작업 모듈은 구현 언어가 자바이기 때문에 코드 상의 통합은 불가능하다. 물론, 자바의 JNI 의 기능을 이용하여 구현할 수도 있겠지만, 그럴 경우 성능도 성능이러니와 자바의 플랫폼 독립성을 상실하게 되므로 고려의 대상에는 포함 시키지 않았다.

영상전화 모듈과 공동작업 모듈은 일단 초기 실행이 되고 난 후에는 매우 밀접하게 연결되어 있을 필요성이 없기 때문에 IPC 를 사용하여 두 모듈간의 통신이 이루어 지도록 한다.

- 공동작업부분

홈서버 중심의 가정내 네트워크 환경에서의 공동작업 서비스를 지원하기 위한 미들웨어로 국제표준 T.120 시리즈를 기반으로 하여 구현하였다. 이는 인터넷의 공동브라우저, 채팅, 화이트보드, 파일전송 등의 기능을 제공한다. 일반전화기의 #버튼을 누르면 실행된다[4].

3.3. SIP 기반의 영상전화 시나리오

SIP 기반의 영상전화 동작 시나리오는 다음과 같다.

- 1) 두 사용자가 일반전화기를 통하여 상대의 ip(상대방 전화번호나 id 번호도 가능)를 눌러서 연결을 시도한다. 상대방이 수화기를 받으면 SIP 기반의 콜 세션이 연결된다. 이로써 음성통화를 할 수 있게 된다.
- 2) 이 때 한 사용자가 필요에 따라 영상통화를 위하여 일반전화기의 *버튼을 누르면 DTV 화면에 상대 및 본인의 모습이 디스플레이 된다.
- 3) 그리고, 필요에 의해 공동작업 모듈을 실행하고자 하면 일반전화기의 #버튼을 누르면 공동작업을 할 수 있는 브라우저가 실행된다.
- 4) 한편, 영상전화의 종료 시에는 위와 비슷한 절차를 밟고 종료한다.

4. 결론

인터넷 멀티미디어 통신 서비스 기술 중 핵심기술의 하나인 VoIP 및 V2oIP 기술은 인터넷과 텔레포니의 통합 네트워크에서 중요한 역할을 하고 있다. 이에 더 나아가 차세대 이동 통신망, 지능망, 인터넷 등이 통합되어 다양한 네트워크 환경으로 가는 추세에 있으며, 이러한 복합적인 네트워크에서 멀티미디어 제어를 위한 새로운 표준 기술로서 SIP 가 연구 개발되고 있다. 특히, 가정내의 네트워크 환경이 급속히 자리잡고 있는 추세에 있고, 이에 정보가전분야가 주목 받고 있다. 이에 SIP 프로토콜 기술은 H.323 에 비해 단순한 구조를 가지고 있고, 새로운 부가 서비

스를 추가하기가 쉽기 때문에 정보가전 등의 많은 분야에서 사용될 것이며, 이미 국내외적으로 다수의 개발 업체들에 의해 개발되고 있는 상태에 있다. 게다가, SIP 기술은 VoIP 또는 모든 호 설정과 접속 설정이 요구되는 모든 서비스(기존의 전화서비스 포함)에 컴포넌트로 사용될 수 있으므로 새로운 부가 서비스 제공 시에 쉽게 장착할 수 있고, 또한 프로그램 크기가 작기 때문에 모든 유무선 단말기에 장착될 수 있어 모빌리티를 제공하여 유무선 인터넷 환경에서 어떠한 비즈니스 모델에도 적용 가능하다고 할 수 있다. 본 논문에서는 SIP 를 기반으로 한 영상전화 시스템에 대해 소개하였다. 이는 임베디드 리눅스인 Qplus 운영체제의 홈서버에 탑재되어 가정내 정보가전의 응용 서비스로 제공되며, 이에 더 나아가 인증, 보안, 감시 기능을 추가하여 개선하고자 한다.

참고문헌

- [1] "SIP: Session Initiation Protocol", draft-ietf-sip-rfc2543bis-o3.ps, internet draft, Internet Engineering Task Force, May 2001
- [2] "SDP: Session Description Protocol", rfc2327, Internet Engineering Task Force, April 1998
- [3] ITU-T Recommendation H.261, "Video Codec for audiovisual Services at Px64kbps," ITU-T, 1994.
- [4] ITU-T Recommendation T.120, "Data Protocols for Multimedia Conferencing," Feb. 1996