

VoiceXML을 이용한 SIP 호 처리 서버의 설계와 구현

김병희^o, 강승찬, 이재오
한국기술 교육대학교 전기전자공학과

Design and Implementation of SIP-Call Processing Server Using VoiceXML

ByoungHee Kim^o, SeungChan Kang, JaeOh Lee
Dept. of Electrical & Electronics Engineering, KUT

요 약

본 논문에서는 SIP(Session Initiation Protocol) 호의 처리 모듈을 가지고 있지 않은 웹 상의 클라이언트들에게 음성서비스 및 멀티미디어 서비스를 제공할 수 있는 SIP 호 처리 서버를 설계 및 구현 하였다. 서버상에서 SIP 호 설정을 위해 사용되는 메시지를 VoiceXML의 태그형태로 변환하는 기능을 제공함으로써 비 SIP 사용자들에게도 SIP 서비스를 할 수 있도록 하였다.

1. 서론

최근 인터넷의 확산 보급과 성능 향상으로 인해 인터넷 사용자가 급증하고 있다. 이러한 영향으로 인터넷 사용자들에 대한 여러 가지의 서비스가 계속해서 요구되고 있다.

그 중 인터넷을 통하여 음성 및 멀티 미디어 서비스를 위한 다양한 통신 프로토콜이 응용 발전 되고 있다. 그 중 활발하게 연구되고 있는 VoIP구현 기술로는 IETF 에서 표준으로 제정한 SIP (Session Initiation Protocol) 와 ITU-T 에서 표준으로 제정한 H.323이 있다.

H. 323은 기존의 음성 데이터를 압축할 수 있는 다양한 G.711, G.723.1, G.728, G.729등과의 인터페이스를 지원하며, 기존의 제품들과의 연계에 있어서 수월하며, RSVP (Resource Reservation Protocol) 등을 이용하

여 인터넷상에서 원하는 QoS를 제공한다는 장점이 있다[3]. 그러나 구조의 복잡성, 새로운 서비스의 확장성, 구현의 어려움과 같은 점들이 문제점으로 대두되면서 텍스트기반의 확장성과 H.323에 비하여 비교적 간단한 구조를 가진 SIP를 이용한 VoIP가 서서히 대두되고 있다[5].

SIP는 세션을 설정하기 위한 절차와 방법을 정하고 있는 프로토콜이다. 호 설정을 위해서는 각 사용자들은 초기의 호 설정을 위한 SIP 메시지를 처리해줄 UA(User Agent)가 필요로 하다. 때문에 SIP처리 모듈이 없는 사용자들은 서비스를 할 수 없는 단점이 있다. 이에 본 논문에서는 SIP 메시지를 처리하여 그 결과값을 웹 상으로 전송 가능한 VoiceXML 문서로 변환하여 처리할 수 있는 서버를 개발하여, 비 SIP 사용자들에게도 SIP 호 설정 서비스를 제공할 수 있도록 하였다.

2. 관련연구

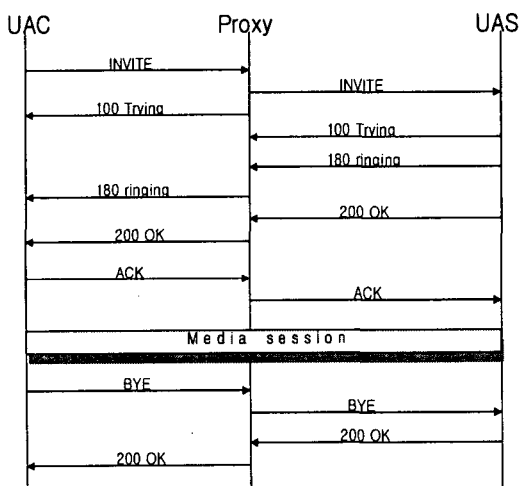
2.1 SIP (Session Initiation Protocol)

SIP는 단말간 또는 사용자들에게 초기 세션의 생성, 수정, 종료를 위한 텍스트 기반 시그널링 프로토콜로, 그 구성은 크게 UA(User Agent) 와 네트워크 서버로 이루어 진다.

UA는 다시 두 가지로 나누어 지는데, 초기 호를 요청하여 세션의 생성을 요구하는 UAC(User Agent Client)와 그 호에 대하여 응답하는 UAS(User Agent Server)로 구성이 된다. 여기서 UAS는 네트워크 서버와는 구분된다.

SIP의 네트워크 서버종류에는 Proxy, Redirect, 등록(Registration)서버가 있다. Proxy 서버는 UA로부터의 요청에 응답 하거나 그 요청을 다음 홉으로 포워딩 한다. Redirect 서버는 UA의 요청에 위치 정보만으로 응답을 하며, 그 요청을 포워딩하는 작업은 하지 않는다. 마지막으로 등록 서버는 UA로부터의 Register 요청을 받아 등록하는 작업을 담당한다.

그림 1 은 SIP의 호 설정의 시나리오를 나타내고 있다.



[그림 1] SIP 기본 호 흐름도

먼저 UA(UAC)가 다른 UA(UAS)에게 호 설정을 위한 Request인 INVITE 메시지를 전송하게 된다. 그러면 Proxy 서버는 이 Request를 UAS에 포워딩 하고

100trying 메시지를 UAC에게 전송하고, 180ringing을 전송하고 대기한다. 얼마 후 UAS의 응답(200OK)을 수신하게 되면 그 응답을, UAC에 포워딩 한다. 마지막으로 UAC로부터의 ACK메시지로 호 설정이 이루어 지게 된다.

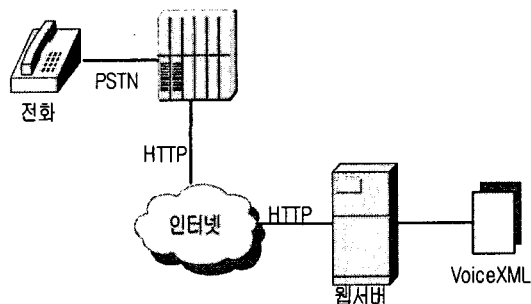
아래의 표1은 SIP 에서 사용되는 각종 Request 메시지와 Response 메시지를 나타내고 있다. [1][2]

[표1] SIP 메시지

Request		Response	
INVITE	콜의 요청	1XX	Informational
ACK	INVITE의 확인응답	2XX	Success
BYE	콜 해지	3XX	Redirection
CANCEL	미결정의 콜 해지	4XX	Client Error
OPTION	Capability 정보요청	5XX	Server Error
REGISTER	위치정보관련	6XX	Global Error

2.2 VoiceXML (Voice eXtensible Markup Language)

VoiceXML은 2001년 3월 당시 AT&T, IBM, Lucent, Motorola 이상 4개사가 주도하여 설립한 민간 포럼인 VoiceXML 포럼에서 제정하여, W3C 가 권고한 마크업 언어로서, 웹 기반에서 음성 인터페이스를 지원하고 개발하기 위한 목적으로 XML을 사용하여 설계되었다. 현재는 VoiceXML 2.0 스펙까지 공개되어 있다.



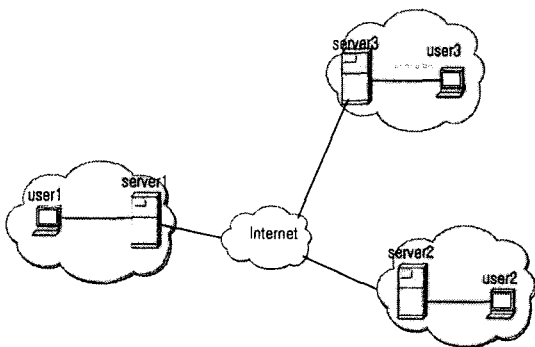
[그림2] VoiceXML의 구성도

그림 2는 VoiceXML의 게이트웨이와 기본동작 시나리오를 나타내고 있다. 그림2에서 PSTN과 웹 사이에서 전화접속기능, VoiceXML의 분석, 렌더링 등의 역할과 웹서버에 VoiceXML 문서를 요청하는 역할을 하는 게이트웨이가 있다. 게이트웨이는 사용자의 요

구에 맞는 서비스를 웹 서버에 요구하고, 다시 사용자에게 응답한다. 웹 서버는 요구되는 서비스의 VoiceXML문서를 웹을 통해 게이트웨이에 전송하게 된다[4][6]..

3. VoiceXML을 이용한 SIP 호 처리 서버

본 논문에서는 이러한 VoiceXML 기술을 응용하여 SIP메시지를 웹 상으로 전송 하고자 한다. SIP 호 처리 서버는 두 가지의 작업으로 구분 될 수 있다. 첫 번째는 서버의 외부에서 입력되는 SIP 메시지를 수신하고, 이를 분석 및 처리하여 헤더 및 바디의 내용들을 분리하게 된다. 분리된 각각의 내용들은 재처리를 거쳐 VoiceXML의 형태로 재구성되는데, 이를 목적지의 클라이언트에게 전송하는 작업이다. 이는 SIP의 처리 모듈을 갖지 못하는 클라이언트에게 세션 초기화를 위한 데이터를 처리 가능한 형태로 전송하는 작업이다. 두 번째는 서버내부의 클라이언트가 보내는 콜 설정을 위한 신호를 수신하고, 이를 분석 및 처리 하여 SIP의 메시지내의 To, From, subject등의 각 필드의 헤더값 등과 SDP(Session Description Protocol)의 바디값으로 재구성한 후, 외부의 목적지로 포워딩을 한 후 그 응답을 기다리게 된다.

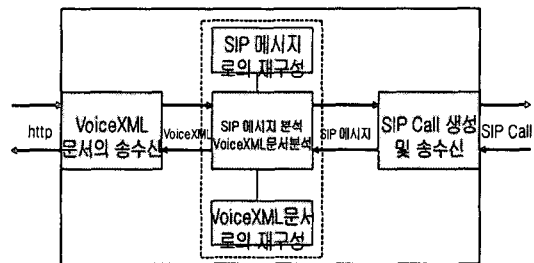


[그림3] SIP 호 처리 서버의 네트워크 구성도

그림 3에서 네트워크 구성을 보여 주고 있다. 세 개의 각 네트워크들은 인터넷을 통해 연결이 되어 있으며, 각각 하나 이상의 클라이언트를 가지고 있다. 각각의 클라이언트인 user1, user2, user3은 모두 SIP처

리 모듈을 가지고 있지 않은 일반의 사용자 이고, 각각의 server1, server2, server3는 모두 SIP 호 신호를 VoiceXML의 형태로 처리 가능한 서버들이다. 동작 형태를 살펴보면, 만약 user1이 user2에 미디어 세션의 설정을 원하게 되면 user1은 server1에 이를 알리게 된다. 그러면 server1은 이 신호를 SIP INVITE 메시지로 변환하여 server2에 호 설정을 요구 하게 된다. server2는 수신하면 server1에 100trying 신호를 전송하고, 이 호 설정을 위한 INVITE신호를 분석하여 INVITE 신호의 To, From, Call-ID, CSeq, Subject, Contact, Type, Length등의 헤더 값들과 바디부분에 표현 되어 있는 SDP의 Media Description, Media Attribute 등의 인자 값을 VoiceXML문서로 재구성한 후, user2에게 전송을 하고 server1측에 180ringing 신호를 전송한다. user2의 수락신호가 수신되면, 이 신호를 수신한 server2는 다시 SIP 200OK 메시지를 생성하여, server1에 전송하게 된다. server1은 이 200OK 메시지를 수신하면 분석하여 다시 VoiceXML의 형태로 재구성 하여 user1에 전송하고 ACK 신호를 기다린다. 마지막으로 user1측에서는 200OK신호를 보내면서 세션을 구성이 완성되게 된다.

그림4는 SIP 호 처리 서버의 구성을 나타내고 있다.



[그림4] 서버구성

①의 부분은 네트워크 내부의 사용자와 VoiceXML 문서를 송수신 하는 부분으로서, 호 설정을 위해 VoiceXML문서로 재 구성된 SIP 메시지를 전송하는 역할을 한다. 또한 SIP 메시지로 재구성 되어야 하는 VoiceXML문서를 HTTP Fetcher를 사용 하여 사용자측에서 받아 오는 작업을 수행한다.

②의 부분은 수신된 SIP 메시지와 VoiceXML문서를

각각 분석 후 재구성하는 부분이다. SIP 메시지는 각각의 필드 값, 바디부분의 인자 값들을 추출하여 VoiceXML의 <assign>, <div>등의 요소 값으로 재구성하는 작업을 수행한다. 반대로 VoiceXML문서들은 XML파서를 이용하여 파싱처리를 한 후, 반환된 데이터들을 SIP 메시지의 필드 값과 바디부분의 내용으로 재구성하는 작업을 수행한다.

③의 부분은 네트워크 외부로부터 호 설정 신호를 수신하고, 또한 새로운 호 설정을 위한 신호를 생성하고, 생성된 설정 신호를 외부로 전송하는 부분이다.

본 논문에서의 구현은 Windows2000서버 환경에서 Visual C++로 구현 되었으며, SIP의 메시지 처리를 위해서는 windows2000서버상에서 Vovida사의 오픈 소스를 사용 하였으며, XML 파서는 마이크로 소프트사의 MSXML 4.0을 사용하였다. SIP의 호 설정 이후 미디어세션은 웨이브파일의 전송으로 시뮬레이션 하였다.

5. 결론

현재 VoIP기술은 음성 데이터를 통합한 네트워크에서 중요한 역할을 하고 있다. 이러한 네트워크에서의 멀티미디어 네트워킹 제어를 위한 표준기술로서 SIP 프로토콜의 연구가 활발히 진행되고 있으며, 특히 SIP 프로토콜은 그 구조가 단순하고 확장이 용이 하기 때문에 정보가전 등의 많은 분야에서 사용될 것이다. 본 논문에서는 SIP 메시지를 각 사용자측에서 처리하는 방식에서 서버가 대신 처리하여 그 결과값만을 http상으로 전송 가능한 VoiceXML 문서로 변환하는 서버를 설계하고 구현함으로써 비 SIP 사용자들에게도 SIP서비스를 제공할 수 있도록 하였다. 향후 영상전화, 영상회의 등의 향상된 서비스를 위해 멀티미디어 데이터의 처리 등의 진보된 성능을 위한 서버의 확장을 위해 연구와 노력이 필요할 것이다.

[참고문헌]

[1] J.Rosenberg and H.Schulzurinne: "SIP: Session Initiation

Protocol", RFC3261, June, 2002

[2] M.Handley and V.Jacobson: "SDP: Session Description Protocol", RFC 2327, April, 1998

[3] ITU-T Recommendation: "Packet based Multimedia Communication Systems", H.323v2 , 1998

[4] VoiceXML Forum: "Voice extensible markup language (VoiceXML) version 1.00",VoiceXML forum specification, VoiceXML Forum, Mar, 2000

[5] Ismail Dalgic, Hanlin Fang: "Comparison of H.323 and SIP for IP Telephony Signaling", 3Com Corporation Technology Development Center

[6] <http://www.voicexml.org>

[7] J.Resenberg and H.Schulzurinne: "Models for Multi Party Conferencing in SIP", draft-ietf-sipping-conferencing-mpdels-01.txt, IETF, July, 2002