Voice over IP – Considerations for a Next Generation Architecture

Markus Hillenbrand, Joachim Götze, and Paul Müller
University of Kaiserslautern
Department of Computer Science
Integrated Communication Systems
67655 Kaiserslautern, Germany
http://www.icsy.de
{hillenbr,j_goetze,pmueller}@informatik.uni-kl.de

Abstract

Voice over IP is on its way to becoming an alternative to the classical telephony system because more and more Voice over IP providers offer their solutions on the Internet and try to increase their clientele. But current Voice over IP technology is still in its commercial and technological infancy and has not yet proven to be worldwide accessible, usable and scalable as the classical telephony system has proven in the past 100 years.

In this paper, an architecture\(^1\) for a next generation Voice over IP model will be outlined and discussed. The main focus lies on interoperability between different Voice over IP providers as well as dependability and robustness.

1 Introduction

The Internet has become one of the most important media for communication. Often the Internet is the basis for a convergence of different media. Today it is expected to use the Internet protocol (IP) not only for the classical Internet services, but also for all types of communication from television to radio and telepho ene. Voice over IP (VoIP), i.e. telephony on the basis of Internet technology, gains more and more importance. Compared with the classical ISDN or PSTN telephony system, VoIP does not only allow a cost saving because of a more efficient use of physical lines in the scope of long-distance calls, but also because of a reduced effort in in-house cabling. Additionally VoIP has potential for providing a new manifold of features beyond telephony [26].

Current products and solutions in the area of Internet-based telephony are highly developed and allow a comparable comfort like their predecessors in the classical telephony, but it is possible to identify several important facts that can be the reason why VoIP still has not yet spread as much as desired by the driving forces:

- Updating VoIP software is very time intensive and often results in high costs for administration while affecting not only clients but also servers as well.
- The simple and user-transparent integration of supplementary services is still not achieved: the introduction of a new supplementary service results in installation and configuration efforts on both client and server side. This is also the case if old hardware has to be replaced by new hardware.
- The support of different heterogeneous clients is at best only possible with different software components that may be incompatible among each other and will only work together with software of the same vendor.
- With regard to the necessity of acquisition and maintenance of additional (and mostly expensive) local hardware, VoIP is difficult to realize especially for small and medium-sized businesses, because the cost-benefit ratio is adverse.

In the Venice project – performed at the University of Kaiserslautern in cooperation with Siemens AG – different architectural approaches for VoIP are under investigation and tested on prototypes. The main focus lies on the previously discussed problem areas and are solved using a service-oriented approach using Web Services and Peer-to-Peer technology. This article provides an overview of this approach and the resulting architectural design. Technical details and related work will not be discussed in detail, they have been published previously and can be gleaned in [20, 37, 38, 39, 18, 19, 16, 17].
Basically, the benefits of the architecture described in this work compared to existing H.323 or SIP-based solutions or products like Skype [36] are the open and service-oriented approach and the migration from client code into distributed services. This makes it easier to develop and integrate new supplementary services into the existing infrastructure and allows for thin clients running on diverse platforms.

In the following, section 2 gives a short introduction into the area of Voice over IP and its related technologies. Section 3 outlines the background of the work, and section 4 explains the architecture developed on an abstract level. Section 5 concludes the paper with a summary and outlook.

2 Voice over IP

With the conflation of the voice network (i.e. ISDN and PSTN) and the data network (i.e. Internet) it is possible to establish an easy-to-use and cheap communication infrastructure. Critics often argue that IP telephony provides less supplementary services than the classical ISDN telephony system. But telephone systems without supplementary services comparable to ISDN are not acceptable in today's business areas. Although many features are not used during the daily business routine, they are still influencing the decision between ISDN and VoIP.

Currently there are two important and standardized protocols for the realization of VoIP: H.323 and the Session Initiation Protocol (SIP). The Multimedia Conference Protocol H.323 [25, 24, 21, 27] of the ITU consists of multiple separate protocols like e.g. H.245 [23] for Control Signaling and H.225 [22] for Call Signaling. H.323 is difficult to implement because of its complexity and a H.323 client application is typically very bulky [32]. On end-user equipment with only few resources like e.g. PDAs it is desirable to have a simple client software saving resources.

An alternative to H.323 is SIP [13, 14, 35, 33] of the IETF. SIP is simpler than H.323 and because of that the client application can be much leaner. SIP is using a human readable protocol (ASCII) instead of a binary signal coding. Both VoIP solutions need to be integrated and administrated in an existing LAN environment using a manifold of software and hardware components like Gateways and Gatekeepers. Hence arbitrary applications have nearly no chance to integrate existing VoIP solutions directly. Furthermore H.323 and SIP both provide only an isolated application: an H.323 client is not capable of communicating directly with a SIP client or vice versa. An H.323/SIP gateway can help to circumvent the problem, but there is no perfect solution for a communication between H.323 and SIP – such an interlinking can only provide basic functionality, because only this functionality exists in both protocols. Most supplementary services cannot be used across these borders [34, 32].

The integration of voice and data communication into a single network allows for the provisioning of a manifold of new supplementary services [31]. Traditional supplementary services like Call Forwarding, Call Waiting, Three-Party, etc. can be extended by the integration of other Internet services like email or instant messaging. A call could be forwarded to a Web page or create a voicemail by using suitable streaming-media tools. Because of these powerful possibilities it is important to constitute a model allowing the fast and easy development and integration of new supplementary services into an existing VoIP system. Many current VoIP products have implemented their supplementary services directly into the server or client application. As a result all clients or server components need a software update in order to enhance or change these supplementary services.

In the next sections a general overview and the architecture of a service-oriented VoIP framework will be discussed. The main focus lies on Call Control (i.e. the means for establishing, managing and disconnecting a phone call and its supplementary services) while Bearer Control (i.e. the management of the actual data transmission over the Internet using QoS) is of less concern.

3 General Overview

In order to circumstantiate the architecture presented in the next few sections, it is necessary to explain some of the major preconditions immanent to the project. There are several hypotheses on which the architecture presented in the next section is built upon. These hypotheses directly lead to some technical and non-technical considerations that have to be dealt with when designing a next generation VoIP infrastructure:

1. All Voice over IP providers act mainly autonomously.

   In the context of a world-wide operable Voice over IP market, this means that every provider is primarily responsible for his own customers. But there also has to be a secure and standardized passage to other providers’ systems. This can also be regarded as a shade of interoperability between all Voice over IP providers.

2. A provider wants to re-use existing infrastructure, i.e. hardware and software.

   Phone companies have invested a lot into their existing hardware and software. Although currently changing backbone infrastructures towards IP technology, they still rely on legacy systems that will prevail for some time. It is thus significant to re-use these components without having to replace them.
3. Supplementary services and their interoperability are the assets of a Voice over IP provider.

Every VoIP provider uses supplementary services to attract customers. This effect has started with the ISDN system and has dramatically increased in mobile communication. Customers are willing to pay for services they can swank with or separate themselves from others. Hence individuality is one key to satisfied customers. Different services at different price levels make sure that every customer can find the combination of services she or he searches for. And the service capabilities offered by one service provider distinguish him from others. Thus, it is a very important issue to be able to integrate and change supplementary services whenever necessary and without great effort. It is imperative that a framework for Voice over IP is designed in that way.

4. Customers are no computer specialists and urge for easy-to-install, easy-to-update, and easy-to-use software.

Current VoIP solutions are mostly realized as flexible software products for personal computers, e.g. [28, 29, 36]. Installation and maintenance is incumbent on the PC’s administrator. But making VoIP accepted on a large scale means to diminish or eliminate administration tasks. Using a VoIP client has to become as easy as using a standard telephone. This can be achieved by creating simple (i.e. less complex) client software and sourcing out to the VoIP provider as much code as possible. The knowledge how to maintain and manage the code is inside the VoIP provider’s domain, and the code should also be located there whenever possible. The clients can then access the code by calling a standardized interface to its functionality.

There are also some constraints that have been taken into consideration:

1. The failure of one component must not result in a complete system failure.

Dependability is one key issue of current software systems and as such it is imperative for an open distributed system to be fault-tolerant. As each component runs on its own server, a failure of such a component (or its communication infrastructure) must not affect the system as a whole. In case of a distributed telephone system, this means that a failing supplementary service (section 4.3) should be compensated for so that a phone call is still possible. And the basic service (section 4.2) responsible for managing a call must always be available. If it fails, a service replica must stand by to take over the old one’s duties.

2. It must be possible to support every established telephony technology without any special treatment by the users.

In the classic telephony system it is possible to make a call to other countries where other telephone standards are in use (e.g. ISDN in Europe and the US). This is transparent to the customer. Telephone providers are responsible for technically accessing other systems. This should also be the case for an Internet based telephony system. Different VoIP providers use different technology (like H.323, SIP, or proprietary products) and should be responsible for translating between these different protocols. Only if that happens transparent to the user, VoIP will prevail.

3. Using existing infrastructure should be easily possible.

Virtualization is the key to leveraging existing technologies or solutions. Building a new system on top of existing systems is the best way to support both older and newer technology at the same time. Telephone providers therefore should have the possibility to extend their existing hardware and software to support a new VoIP infrastructure.

In order to encompass all these considerations, Web Service technology [4, 15, 12, 2, 6, 5, 11, 1, 3, 7] has been chosen to become the realization platform for the project’s prototypes. Web Services technology can be used to leverage legacy systems and offer their functionality in a service-oriented environment [10, 9]. And it has proven to be an acceptable means for making these legacy systems accessible by current (i.e. newer) technology while allowing for innovative solutions by abstracting from an underlying concrete implementation or technology.

4 Architectural Overview

The services offered by a VoIP provider can be classified into three distinct categories: Management Services are services necessary for the common usage and maintenance of services. They are not bound to VoIP services and can be used in any service-oriented application scenario. The Basic VoIP Services contain all elementary functions that are responsible for making a phone call with other participants. The category of Supplementary Services finally contains all additional services that complement the Basic VoIP Services and contribute comfort before, during or (shortly) after making a phone call. An overview of this architecture can be seen in figure 1 and will be discussed in the next subsections.
4.1 Management Services

Several Management Services help VoIP providers to actually create and maintain a distributed and open Voice over IP infrastructure. Currently, the following services can be identified:

- **Role-based Single Sign-on (SSO).** In order to access all services of a VoIP provider, a token-based single sign-on strategy is the most promising approach [8, 16, 17]. A user has to authenticate himself once and will then receive a token allowing him to proof his identity in any further communication by providing this token. If the user is accessing a service, this token will allow the user not only to proof his identity, but also to proof his right to access this service. As there may be several services involved in the process of initiating, making, or terminating a phone call, it is very important to use an authentication and authorization strategy based on single sign-on. This strategy allows the user to be authenticated to any service inside the authentication domain without the need to enter his credentials more than once. Additionally, if a user can act under several roles, it is imperative to allow for changing his role during service access with respect to priority and accounting issues. This might be the case if a user is working in several projects utilizing restricted services. Beyond the scope of a single VoIP provider, the interaction of users from many different VoIP provider has to be considered, too. In order to support this dynamic interaction of users the VoIP providers themselves form a federation of trust. This federation of trust allows a user to initiate and receive calls or change the settings of the supplementary services he is using while being connected with a VoIP provider that is not the VoIP provider the user is articulated to by a contract. A more detailed description of this token-based single sign-on solution developed within the project can be found in [17, 16].

- **Metering, Accounting, and Billing (MAB).** Because of the service-oriented approach used in this architecture, it is possible to use services that provide the same service with different characteristics, e.g. quality of service. All these services need not to be made available by a single provider, but it is possible to use services of different providers. As a result a multitude of different service providers can be involved in a single phone call. In order to have a reasonable accounting for services used, the VoIP providers need to have a metering system that keeps track of all requested and utilized services in the VoIP system.

- **Software Deployment (SDS).** In order to provide an easy-to-use VoIP application to the end-user, it is im-
portant to dispense the user from tasks like installing and/or updating software. A Software Deployment Service allows the user to use an up to date application without any further effort. Additionally this service enables a VoIP provider not only to maintain a state of the art software infrastructure, but also allows the provider to automatically and dynamically replicate certain services in order to compensate high load situations and establish a replica management. As a result a Software Deployment Service reduces the administration effort on both sides, user and service provider. A comprehensive overview of the Software Deployment Service developed can be found in [18].

- **Information Brokering (IB).** Due to the potentially huge amount of diverse information, it is important to have a service that is responsible for brokering, searching, and obtaining desired information, while also allowing to scale up to a growing quantity if for example new services are introduced by service providers. In order to support such a scalability, information brokering should be based on a Peer-to-Peer (P2P) network, which is capable of handling and managing large amounts of network nodes providing multiple information types. Because of the stability, in terms of propagating advertisements of information, a P2P network can provide a reasonable advantage compared to other architectures, e.g. a client/server-architecture like UDDI [30]. In this context information is used as a synonym for services (including VoIP providers and supplementary services) and client information. These information allow a potential service consumer to find a suitable service provider, e.g. information about how to connect to a selected supplementary service or which VoIP provider a certain customer is currently using. The latter is especially important for routing a phone call to the customer’s current provider.

- **Feature Interaction Manager (FIM).** This manager component is commissioned to check any request for a supplementary service of the caller or the callee with respect to their usability and applicability. The main focus is to ensure that there is no undesired side-effect because of the consecutive use of multiple supplementary services. Any recognized side-effect is being detected and resolved by using a knowledge base to provide the maximal functionality of the supplementary services requested.

As all of these components are realized as Web Services, their interfaces [20, 18, 16] have to be well designed and agreed-upon. Only this makes interoperability between service providers possible and allows the customer of one provider to talk to a customer of another domain.

### 4.2 Basic VoIP Services

The actual VoIP Service is the core component of the architecture. This component provides the basic functionality for Call Control, i.e. allows a client to initiate, perform and terminate a phone call. This implies that this component is directly connected to the client application and provides abstraction from the underlying telephony technology. As a result it is possible to mediate a communication that has been initiated by a SIP/H.323 client or that is targeted on a SIP/H.323 client without the need of interaction with the user or that this mediation becomes recognizable by the user [37, 38]. A special component is responsible for finding VoIP users in other domains and routing the phone call through the system to the other provider’s VoIP service (i.e. WSBridge in figure 1). Additionally the VoIP Service has to supervise the communication, so that it is possible to redirect media streams to other endpoints, e.g. if an user has been set on hold, while the other participant of the phone call is receiving a second call. Directly coupled with this basic service is the Feature Interaction Manager.

This also means that this component always holds a state for every phone call currently running in the system. With respect to the supplementary services, this denotes that the basic VoIP Service has to offer an interface to manipulate this state. This interface should never be exposed to the outside world and be accessible by the VoIP providers services only. This is assured by using the role-based single sign-on service.

### 4.3 Supplementary Services

Supplementary services are placed around the core VoIP Service and make it worthwhile for customers to use a specific VoIP provider. They offer functionality that can be combined until a customer’s needs are satisfied. Typical supplementary services for private and business customers are Call Waiting, Call Forwarding, Call Hold, Three-Party, etc. And there are extensive supplementary services like a call center that can be found in larger companies. Currently, there are several hundred supplementary services available in ISDN and GSM networks – some more, some less known [26, 21].

#### Integrating Supplementary Services

It is imperative for a VoIP provider to be able to integrate new supplementary services easily into a VoIP framework. A new or updated feature should not always result in code changes inside other components of a distributed system. It is therefore necessary to provide the means for an easy integration of supplementary services into a VoIP framework. Figure 2 illustrates the solution proposed here and
Figure 2. Supplementary Services for VoIP

shows one exemplary supplementary service inside a VoIP provider’s domain.

Every service available in the VoIP framework has three different access points (currently realized as service ports in Web Services terminology) where service requestors may perform a method call. The public part can be accessed by any participant of the VoIP framework (e.g. role=user in figure 2). Its interface is published once and should remain stable as this part is accessed by all VoIP clients. The second part of the service’s functionality is only accessible by other services of the VoIP system (role=voip in figure 2). Here several internal management and service functions are appointed to be used by internal (i.e. with respect to the VoIP provider) entities. The third category contains all administration functionality only accessible by the VoIP provider’s administration role users (role=admin in figure 2). Here all service internals can be managed. Together with the SSO service this allows for a fine-grained access control to the Web service functionality.

Accessing Supplementary Services

The internal interface of the basic VoIP Service is the key to manipulating call related data on the client side (i.e. state information like IP ports, RTP parameters, etc.). It offers functionality to the other internal supplementary services that might want to manipulate a phone call session either before, during, or shortly after an actual call. All the messages meant for one of the participants have to be initiated by a supplementary service and are actually issued by the basic VoIP Service. Thus, the clients have to accept incoming (i.e. not actively requested) messages from the basic VoIP Service only. But the clients can actively invoke a supplementary service and directly obtain the result. If a service invocation results in manipulating the internal state of a call, it will again be managed by the basic VoIP Service.

The functionality of the supplementary services are described and published using WSDL [6], so that clients and other services can access the supplementary services using an appropriate binding. The Single Sign-on Service (SSO) is then used to guarantee access to functions only to those users and services acting in the necessary role – others won’t obtain a suitable token for the service from the SSO. Interoperation between different VoIP providers is assured by using well-defined XML data types (defined with XML schemas) and well-defined WSDL interface definitions for each and every service inside the VoIP framework.

Consuming Supplementary Services

The usage of supplementary services can be initiated by either a VoIP client or a VoIP service running in the system. A client can always ask the Information Broker of his VoIP provider to obtain a list of all supplementary services that are usable in the current state of the client. Of course the Information Broker will take into account the services a client has booked with the VoIP provider and ignore any additional service. The current state of the client determines which supplementary service a client can potentially use, e.g. a Call Transfer service can only be invoked when a phone call has already been established. Additionally, the Information Broker tells the client if the supplementary service provides any additional software component that should be integrated into the client’s user interface. This software component will then be downloaded using the Software Deployment Service (SDS) and integrated into the client. As such the client software adapts to the needs of the client and offers only the services usable in a current state. This mechanism makes client software lean and ease-to-use.

4.4 Example

The example in figure 3 explains the process of using and administrating two supplementary services with three distinct clients.

Client 1 wants to call Client 2, but Client 2 has already activated and configured the supplementary service Call Forwarding. This supplementary service allows Client 2 to redirect his incoming phone calls to another VoIP client, i.e. Client 3. Additionally all clients are using a different VoIP provider and every intermediate VoIP provider has to check which VoIP provider is used by the client addressed. While
Client 2 has redirected all incoming calls to Client 3, the call of Client 1 is reaching Client 3, which itself is currently offline. So Client 1 has to leave a message on the Answering Machine of Client 3, which is also a supplementary service of the VoIP provider of Client 3 and that has already been activated and configured by Client 3 prior to the incoming call.

5 Summary and Outlook

In this paper, an architecture for a next generation Voice over IP framework has been outlined and discussed. The main focus of the architecture is on interoperability between different Voice over IP providers as well as dependability and robustness.

In the beginning the actual situation on the VoIP market has been described, including an overview of the two important communication standards for VoIP: H.323 and SIP. Although not many supplementary services are used on a regular basis during common telephone communication, the possibilities provided by the supplementary services based on Internet technology are much more extensive than in classical telephone communication. As a result the supplementary services have been an additional focus of the architecture described here.

The requirements of a suitable VoIP architecture are the basis of this approach. Starting with these requirements the different components of the architecture have been explained: Management Services, Basic VoIP Services, and Supplementary Services. While the Basic VoIP Services provide the basic VoIP functionality of the architecture, the Management Services provide functionality like authentication and authorization, information brokering, metering and accounting, or feature interaction management. Especially feature interaction management is highly important, because of the rapid growth of supplementary services that can be used. Feature interaction is a result of the consecutive usage of multiple supplementary services, that has to be addressed to provide a maximum of the possible outcome the supplementary services can offer.

As the process of developing a prototype for the architecture described here comes to an end, the future work is dedicated to extensive tests with the focus on scalability, reliability and dependability of the prototype. These tests will prove that the requirements expressed in the beginning are met. Additionally the quality of the feature interaction management will be checked thoroughly in order to provide a large amount of possibilities to use supplementary services consecutively.

References
