

Desenvolvimento de um Software Livre de Videophone baseado em VoIP

Yuri C. Gomes, Diego M. O. Bezerra, Fernanda B. C. Campos,
Joelma B. Silva, Danilo F. S. Santos, Hyggo O. Almeida e Angelo Perkusich *

¹Laboratório de Sistemas Embarcados e Computação Pervasiva
Centro de Engenharia Elétrica e Informática
Universidade Federal de Campina Grande - UFCG
C.P. 10105, 58.109-970 – Campina Grande – PB – Brasil

{yuricgomes,danilo,hyggo,perkusic}@dee.ufcg.edu.br,

{diego.bezerra,joelmabsilva}@gmail.com, fernanda_baracuy@yahoo.com.br

Abstract. *This paper presents a VoIP-based infrastructure for a videophone application. We describe its basic architecture and the communication protocols involved in the aggregation of video-conference resources to construct a videophone application. We present the adaptations performed in the Shtoom soft-phone and tools used for developing the videophone application.*

Resumo. *Este trabalho apresenta, no contexto de VoIP, uma infraestrutura para uma aplicação de videophone. Descrevem-se a arquitetura básica e os protocolos de comunicação envolvidos para agregar o recurso de videoconferência e obter uma aplicação videophone. Apresentam-se as adaptações realizadas no softphone Shtoom e as ferramentas utilizadas no desenvolvimento do aplicativo.*

1. Introdução

A sempre crescente necessidade de obter informação e compartilhá-la, decorrente da globalização que aumenta a interdependência mundial em todos os aspectos, ganhou enorme força com o surgimento da rede mundial de computadores, a Internet. A partir da popularização desta, infra-estruturas de comunicação atingem taxas de crescimento surpreendentes. Por todo o mundo há grandes investimentos na área de comunicação, sintonizados com a quantidade de serviços disponíveis, bem como com a demanda do mercado consumidor.

Atualmente, uma grande revolução vem acontecendo nas telecomunicações, resultante do incrível crescimento das redes baseadas em pacotes, especialmente da consolidada Internet que tornou-se o grande meio de comunicação mundial [Tanenbaum 2003].

Esta grande revolução refere-se à tecnologia VoIP (*Voice Over Internet Protocol*) que consiste no tráfego de voz utilizando uma infra-estrutura de rede de pacotes usando o protocolo IP (*Internet Protocol*), obtendo, assim, uma só rede convergente [Khasnabish 2003]. Esta tendência não indica apenas um movimento de mercado, mas demonstra o progresso da tecnologia de rede. A telefonia baseada em redes de pacotes

*Os autores agradecem o apoio da Nokia do Brasil e Instituto Nokia de Tecnologia pelo apoio ao desenvolvimento deste trabalho.

encontra-se em crescente maturidade e muitos serviços estão sendo implantados neste contexto.

Uma vantagem fundamental de VoIP é a possibilidade da substituição do telefone tradicional por uma aplicação de computador conhecida como *softphone*. Trata-se de um telefone IP implementado em *software* [Brandl et al. 2004]. As aplicações de um *softphone* podem ser adaptadas para oferecer serviços multimídia, possibilitando não somente uma chamada telefônica, mas também a realização de reuniões virtuais, tornando a comunicação mais interativa pelo uso de videoconferência.

Este trabalho descreve a adaptação de um *software* livre de *softphone* para a criação de um *videophone*. São detalhados os princípios básicos envolvidos neste tipo de comunicação, como também as alterações necessárias e as ferramentas de software livre utilizadas.

2. Arquitetura Básica

Conceitualmente, a grande diferença entre VoIP e a telefonia convencional de voz está na maneira como a voz é transmitida através do meio [Khasnabish 2003]. A tecnologia VoIP é baseada em padrões e recomendações e aprovado por institutos de padronização internacional, como por exemplo, o IETF [Rosenberg et al. 2002] ou o ITU-T [H.323 1999]. Ainda não há um único e definitivo padrão formal sobre como implementar VoIP, e sendo assim, cada fabricante possui seu próprio padrão. Há, portanto, várias opções sobre como utilizar o IP para transportar pacotes de voz.

No entanto, a rede de dados não foi projetada para transportar voz. Isso significa que não há garantias de entrega dos pacotes em tempo real para o protocolo IP. Essa é a maior deficiência dos protocolos de internet: a sua incapacidade em transportar voz e vídeo em tempo real. Os níveis de protocolos superiores devem, então, implementar mecanismos para garantir este requisito. Como solução, diversas especificações de protocolos completos para transmissão de várias formas de dados multimídia em tempo real foram criados.

No contexto de telefonia IP, dois padrões principais disputam a hegemonia: o H.323, [H.323 1999], proposto pela ITU-T, presente em muitos equipamentos e aplicativos de *software* VoIP; e o SIP (*Session Initiation Protocol*) [Rosenberg et al. 2002], proposto pela IETF, que, apesar do curto tempo do processo de padronização, tem mobilizado muitos fabricantes da área de telefonia e dados, em virtude da sua flexibilidade, aderência a padrões genuinamente Internet e de arquitetura aberta.

O protocolo de sinalização adotado em nossa arquitetura proposta foi o SIP, devido principalmente ao fato de apresentar uma enorme facilidade de implementação e de adição de novos serviços, tais como mensagens instantâneas, video conferência, entre outros.

O protocolo SIP utiliza o protocolo de descrição de sessão, SDP (*Session Description Protocol*) [Handley and Jacobson 1998], para informar a descrição de parâmetros da sessão entre os agentes. Este protocolo é encapsulado na mensagem INVITE do SIP em um formato de texto simples. Com isso, os agentes podem compartilhar informações sobre a mídia antes de ingressar em uma sessão. O SIP não é um sistema de comunicação verticalmente integrado, pelo contrário, é um componente que pode ser usado com outros protocolos IETF para construir uma arquitetura multimídia completa.

A transmissão tradicional de dados prima pela confiabilidade entre os pontos-fins, fazendo uso de recursos como controle de erro, retransmissão de pacotes, etc. No entanto, a transmissão de dados em tempo real (transmissão de voz e/ou vídeo), não necessita de tal confiabilidade visto que um dado retransmitido pode ser mais danoso ao desempenho do sistema do que um dado perdido. Assim, redes de voz sobre IP devem deixar que o controle de erro seja feito pelos níveis mais elevados.

Como o serviço fornecido pelo protocolo IP é sem conexão, cada datagrama é tratado como uma unidade independente, e a comunicação é não confiável, pois não são utilizados reconhecimentos fim a fim ou entre nós intermediários. Em geral, existem dois protocolos disponíveis na camada de transporte da arquitetura TCP/IP quando transmitindo informação através da rede IP. Estes são o TCP e o UDP.

O protocolo de transporte utilizado para transmissão em tempo real é o UDP, visto que ele apresenta uma sobrecarga (tempo de latência da rede) menor que o TCP, e em virtude do caráter da aplicação que é em tempo real.

Aplicações para transmissão de voz têm restrições de tempo real, e mecanismos devem ser oferecidos para garantir o recebimento da informação na seqüência correta, com boa confiabilidade e com características de atraso previsíveis. Para isso, o protocolo RTP (*Real-Time Transport Protocol*) [Schulzrinne et al. 2003], é usado no transporte de mídia (áudio e/ou vídeo). Isto garante que todos os pacotes recebidos serão processados na mesma ordem de envio e possibilita, através de informações provenientes de seu cabeçalho, realizar transmissão de dados em tempo real sobre a rede de comutação de pacotes.

No entanto, o protocolo RTP não reserva recursos da rede, apenas fornece informações a fim de que o receptor possa compensar perdas e atrasos de pacotes. Um protocolo para esta finalidade, que utiliza a mesma pilha e trabalha conjuntamente com o RTP é o RTCP (*Real-Time Control Protocol*) [Schulzrinne et al. 2003]. Ele é utilizado para realizar o controle da qualidade de serviço da seção ativa.

A arquitetura básica proposta para uma aplicação que utiliza SIP, que é ilustrada na Figura 1, foi definida, fundamentalmente, segundo as necessidades do tipo da aplicação. Observa-se na Figura 1 a existência de um fluxo de pacotes SIP, responsável pela iniciação da comunicação, e um fluxo de mídia, que transporta as informações de áudio e vídeo, utilizando os protocolos RTP e RTCP. As camadas inferiores dessa arquitetura (UDP e IP) transportam ambos os fluxos.

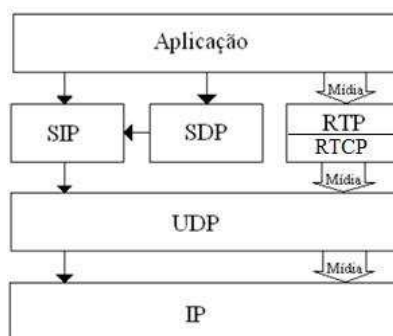


Figura 1. Pilha de protocolos para uma aplicação VoIP utilizando SIP

3. Ferramentas de *software* livre utilizadas

Várias ferramentas de *software* livre foram utilizadas. Um *softphone* que é base da arquitetura proposta é o Shtoom [Shtoom 2004], desenvolvido por Anthony Baxter, de código aberto sob a licença GNU LGPL, implementado em Python [Python.org 2005]. O Shtoom, versão 0.2, é portátil para as plataformas Windows, Linux/Unix e Mac OS X e disponibiliza algumas opções de interface gráfica como GTK/GNOME, Tk/Tkinter e linha de comando. Ele também tem suporte de áudio OSS (*Open Sound System*) em Linux.

No entanto, esta aplicação requer a instalação do interpretador Python, versão 2.3 ou superior, e também do *framework* de rede Twisted [Fettig 2005], versão 1.3 ou superior, utilizado para o desenvolvimento de aplicações de rede, que inclui implementações de um número de serviços de rede frequentemente usados, tais como servidor *web*, *chat*, *mail server*, entre outros.

Para o entendimento do Shtoom, foi de fundamental importância a utilização de três ferramentas: O Epydoc [Epydoc 2005], um gerador de documentação para Python; o DIA [Dia 2005], um gerador de UML; e o pyDEV [PyDev 2005], um plugin para a plataforma Eclipse, fornecendo suporte para desenvolvimento em Python e permitindo utilizar a ferramenta de depuração disponível no referido IDE.

Mais uma ferramenta utilizada neste trabalho foi o GIMP *toolkit*, abreviação GTK+, que é uma popular ferramenta de *widgets* para sistemas de janelas X. GTK foi utilizado para criar a interface gráfica do usuário. Apesar de ser escrito em C, existem *bindings* para a construção de programas em C++, Perl, Java, Python e PHP. No caso de Python, a interface PyGtk [PyGTK 2005] foi utilizada para usar os recursos da biblioteca GTK.

4. Adaptação do *softphone* para a criação do *videophone*

O propósito principal deste trabalho consiste na adição do recurso de videoconferência à uma aplicação *softphone*, no caso o Shtoom. Para a adaptação deste *softphone* para um *videophone*, foram feitas algumas modificações em alguns módulos.

Primeiramente, foi alterado o módulo SDP, que é a implementação do protocolo SDP, para que este pudesse descrever uma sessão que comportasse vídeo. Para isso, foi acrescentado à mensagem SDP uma linha contendo vídeo como tipo de mídia transportado. Com esta modificação o servidor SIP reserva duas portas para a comunicação, uma para áudio e outra para vídeo.

Além de modificar a mensagem SDP, também foi necessária a inclusão de uma implementação do protocolo RTP para o envio de imagens. Em outras palavras, criou-se um módulo a mais, em Python, contendo o código necessário para que os pacotes com as informações de imagens pudessem ser transportados segundo o protocolo RTP.

O processo de captura e transmissão das imagens consiste no envio das imagens obtidas por uma WebCam USB QuickCam Messenger da Logitech. Neste processo, o tamanho da imagem é verificado para determinar se há necessidade de fragmentação da informação em pacotes, ou não. Em seguida, os pacotes são encapsulados com cabeçalho RTP e transmitidos. O receptor recebe os dados e confere, através dos campos contidos

no cabeçalho RTP, a validade da informação, repetindo-se o procedimento até que haja uma interrupção da transmissão por um dos clientes.

Com relação ao controle de requisição e envio de vídeo foram criados pacotes especiais que contêm informações importantes no cabeçalho, como campos de requisição e aceitação. Com as informações definidas no cabeçalho do pacote, o usuário tem a possibilidade de enviar e responder a requisições de vídeo, optando pelo recebimento de vídeo, ou não.

Na Figura 2 é apresentada a tela principal de interface gráfica com o usuário do *videophone*. Já na Figura 3 é apresentada a tela para configuração com os dados do usuário, de forma a registrar-se junto ao servidor SIP disponível.



Figura 2. Tela principal do MORAL Videophone



Figura 3. Tela de configuração do usuário

5. Conclusão

O protocolo de iniciação de sessão - SIP, vem sendo cada vez mais utilizado, tornando-se um protocolo de sinalização padrão para estabelecimento de comunicação fim-a-fim entre diferentes dispositivos usando diferentes plataformas. Entre tais dispositivos podem-se citar computadores pessoais, telefones IP, aparelhos celulares, *handhelds*, entre outros. Dessa forma, os desenvolvedores de aplicativos neste contexto devem estar familiarizados com as especificações deste protocolo e com suas implementações e ferramentas de suporte, visto que uma comunicação fim-a-fim será um requisito essencial em novas aplicações.

A estrutura básica proposta apresenta-se bastante simplificada proporcionando boa facilidade de implementação, sendo validada através do *videophone*, obtido a partir

da aplicativo Shtoom, com desempenho bastante satisfatório. Vale ressaltar a importância fundamental das ferramentas de *software* livre utilizadas durante o desenvolvimento do *videophone*, que foram determinantes na viabilização da implementação do aplicativo.

Por fim, a contribuição das comunidades de *software* livre, tanto nacionais quanto internacionais, proporcionou um intercâmbio de conhecimento promissor que pode ser utilizado para o desenvolvimento de projetos muito maiores do que a aplicação de *videophone* apresentada neste artigo.

Referências

- Brandl, M., Franzens, K., Daskopoulos, D., Dobbelsteijn, E., Giuseppe Garroppo, R., Janak, J., Kuthan, J., Niccolini, S., Ott, J., Prella, S., Ubik, S., and Verharen, E. (2004). *IP Telephony Cookbook*. John Wiley e Sons Inc.
- Dia (2005). <http://www.gnome.org/projects/dia/links.html>.
- Epydoc (2005). Epydoc Release 2.1. <http://epydoc.sourceforge.net/>.
- Fettig, A. (2005). *Twisted Network Programming Essentials*. em <http://twistedmatrix.com/>.
- H.323 (1999). H.323 Recommendation, Packet-Based Multimedia Communications Systems. International Telecommunications Union (ITU-T).
- Handley, M. and Jacobson, V. (1998). SDP: Session Description Protocol. RFC 2327, Internet Engineering Task Force. <http://www.ietf.org>.
- Khasnabish, B. (2003). *Implementing Voice over IP*. John Wiley e Sons Inc.
- PyDev (2005). Python IDE plugin for Eclipse. <http://sourceforge.net/projects/pydev/>.
- PyGTK (2005). PyGTK: GTK+ for Python. <http://www.pygtk.org/>.
- Python.org (2005). <http://python.org/>.
- Rosenberg, J., Schulzrinne, H., Schooler, E., Handley, M., Camarillo, G., Johnston, A., Peterson, J., and Sparks, R. (2002). SIP: Session Initiation Protocol. RFC 3261, Internet Engineering Task Force .
- Schulzrinne, H., Casner, S., Frederick, R., and Jacobson, V. (2003). RTP: A Transport Protocol for Real-Time Applications. RFC 3550, Internet Engineering Task Force.
- Shtoom (2004). Shtoom Release 0.2. <http://www.divmod.org/projects/shtoom>.
- Tanenbaum, A. S. (2003). *Computer Networks*. Prentice-Hall, 4th edition.