

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE CIÊNCIA DA COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO



**MODELAGEM E SIMULAÇÃO ESTOCÁSTICA APLICADA AO
PLANEJAMENTO DE CAPACIDADE DE SISTEMAS GATEWAY DE
VOZ**

MURIEL RIBEIRO ALVES

Uberlândia, Minas Gerais

2014

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE CIÊNCIA DA COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO



**MODELAGEM E SIMULAÇÃO ESTOCÁSTICA APLICADA AO
PLANEJAMENTO DE CAPACIDADE DE SISTEMAS GATEWAY DE
VOZ**

MURIEL RIBEIRO ALVES

Dissertação de Mestrado apresentada à Faculdade de
Ciência da Computação da Universidade Federal de
Uberlândia, Minas Gerais, como parte dos requisitos
exigidos para obtenção do título de Mestre em Ciência
da Computação.

Área de concentração: Engenharia de Software
Orientador: Prof. Dr. Rivalino Matias Jr.

Uberlândia, Minas Gerais

2014

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

A474m Alves, Muriel Ribeiro, 1981-
2014 Modelagem e simulação estocástica aplicada ao planejamento de
capacidade de sistemas gateway de voz / Muriel Ribeiro Alves. - 2014.
81 f. : il.

Orientador: Rivalino Matias Jr.
Dissertação (mestrado) - Universidade Federal de Uberlândia,
Programa de Pós-Graduação em Ciência da Computação.
Inclui bibliografia.

1. Computação - Teses. 2. Simulação (Computadores) - Teses. I.
Matias Júnior, Rivalino. II. Universidade Federal de Uberlândia.
Programa de Pós-Graduação em Ciência da Computação. III. Título.

CDU: 681.3

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE CIÊNCIA DA COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Os abaixo assinados, por meio deste, certificam que leram e recomendam para a Faculdade de Computação a aceitação da dissertação intitulada "**Modelagem e Simulação Estocástica Aplicada ao Planejamento de Capacidade de Sistemas Gateway de Voz**" por **Muriel Ribeiro Alves** como parte dos requisitos exigidos para a obtenção do título de **Mestre em Ciência da Computação**.

Uberlândia, 29 de Agosto de 2014

Orientador: _____

Prof. Dr. Rivalino Matias Jr.
Universidade Federal de Uberlândia

Banca Examinadora:

Prof. Dr. Lúcio Borges de Araújo
Universidade Federal de Uberlândia

Prof. Dr. Paulo José de Freitas Filho
Universidade Federal de Santa Catarina

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE CIÊNCIA DA COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Data: Agosto/2014

Autor: **Muriel Ribeiro Alves**
Título: **Modelagem e Simulação Estocástica Aplicada ao Planejamento de Capacidade de Sistemas Gateway de Voz**
Faculdade: **Faculdade de Computação**
Grau: **Mestrado**

Fica garantido à Universidade Federal de Uberlândia o direito de circulação e impressão de cópias deste documento para propósitos exclusivamente acadêmicos, desde que o autor seja devidamente informado.

Autor

O AUTOR RESERVA PARA SI QUALQUER OUTRO DIREITO DE PUBLICAÇÃO DESTE DOCUMENTO, NÃO PODENDO O MESMO SER IMPRESSO OU REPRODUZIDO, SEJA NA TOTALIDADE OU EM PARTES, SEM A PERMISSÃO ESCRITA DO AUTOR.

Dedicatória

Aos meus pais, Jair e Suelene.

Ao meu irmão Henry.

À minha irmã Luana.

Agradecimentos

Ao Professor Rivalino,
meu orientador, pela sua paciência, apoio e conselhos.

RESUMO

Este trabalho apresenta uma abordagem baseada em modelagem estocástica e simulação discreta para realizar o planejamento de capacidade de sistemas *gateway* de voz. A proposta foi avaliada por meio de um estudo de caso, envolvendo a análise e planejamento da capacidade de um sistema *gateway* de voz real utilizado em uma IFES, o qual processa uma demanda média superior a 1000 chamadas diárias. Por serem sistemas de natureza crítica para as organizações, com possibilidade de diferentes entroncamentos para diferentes operadoras de telefonia e possuindo vários tipos de recursos necessários para completar e atender chamadas simultaneamente, o nível de complexidade desses sistemas torna proibitivo o uso de modelos analíticos para o seu planejamento de capacidade. Além da flexibilidade em modelar sistemas com elevada complexidade, observou-se também que os recursos de visualização da simulação computacional se mostraram importantes durante a etapa de elicitação de conhecimento dos técnicos responsáveis pelo sistema investigado, principalmente para a validação do modelo proposto. A parametrização do modelo de simulação foi realizada por meio da caracterização estatística dos dados reais dos sistemas PABX e *gateway* de voz analisados. Para tanto, foi necessário o uso de métodos estatísticos paramétricos e não paramétricos. Entre os resultados obtidos, a abordagem proposta permitiu, durante o estudo de caso, identificar alternativas de configurações que otimizaram o uso de recursos do *gateway* de voz, em especial do seu entroncamento E1, aumentando assim sua disponibilidade em 14% sem a adição de novos recursos. Isso representou um acréscimo de 2926 novas chamadas diárias sem novos investimentos.

Palavras-chave: Simulação Discreta; Planejamento de Capacidade; Gateway de Voz.

ABSTRACT

This work presents a systematic approach, based on stochastic modeling and discrete-event simulation, to plan the capacity of voice gateway systems. The proposal was evaluated through a case study, related to the analysis and capacity planning of a real voice gateway system from a Federal University. The modeled system processes, in average, more than 1000 phone calls daily. Given that voice gateways are critical to organizations, connecting to different providers, and requiring different resources to complete simultaneous calls, the use of analytical models are not adequate to plan and analyze their capacity. In addition to the flexibility to deal with high complexity modeling scenarios, the visualization of computer simulation was important during the engineers' knowledge eliciting, mainly in the model validation step. The simulation model parameterization was conducted through the characterization of real call record data from the PABX and voice gateway analyzed. For this purpose, it was necessary the use of parametrical and non-parametrical statistical methods. Based on the proposed approach, it was possible to identify alternative setup for the study case, which allowed optimizing the use of the voice gateway's resources, especially its E1 trunk, increasing its availability in 14% with no addition resources. This represents an increase of 2926 new daily calls without any investment.

Keywords: Discrete Simulation; Capacity Planning; Voice Gateway.

SUMÁRIO

LISTA DE TABELAS.....	11
LISTA DE FIGURAS.....	12
LISTA DE ABREVIATURAS E SIGLAS	13
CAPÍTULO 1 – INTRODUÇÃO	15
1.1 – Contextualização.....	15
1.2 – Relevância do Trabalho	16
1.3 – Objetivos da Pesquisa	16
1.3.1 – Objetivo Geral	16
1.3.2 – Objetivos Específicos	16
1.4 – Desenvolvimento da Pesquisa.....	17
1.4.1 – Revisão da Literatura.....	17
1.4.2 – Definição do Estudo de Caso.....	17
1.4.3 – Coleta de Dados.....	18
1.4.4 – Análise dos Dados de Entrada.....	18
1.4.5 – Criação e Validação do Modelo de Simulação.....	18
1.4.6 – Execução das Simulações	18
1.4.7 – Análise dos Resultados	18
1.5 – Estrutura do Trabalho	19
CAPÍTULO 2 – REVISÃO DA LITERATURA	20
2.1 – Introdução	20
2.2 – Modelagem e Simulação.....	20
2.2.1 – Simulação Discreta de Eventos	21
2.2.1.1 – Mecanismos de Avanço do Tempo	22
2.2.1.2 – Números Aleatórios e Variáveis Aleatórias	22
2.2.2 – Linguagens de Simulação	24
2.2.3 – Testes de Aderência.....	25
2.2.3.1 – Kolmogorov-Smirnov	26
2.2.3.2 – Anderson-Darling.....	26
2.2.3.3 – Qui-Quadrado.....	27
2.2.4 – Gráficos de Probabilidade	28
2.2.5 – Intervalo de Confiança para a Média.....	28

2.3 – Protocolos VoIP	29
2.3.1 – SIP	29
2.3.2 – RTP.....	32
2.4 – Trabalhos Correlatos.....	32
CAPÍTULO 3 – SISTEMA MODELADO NO ESTUDO DE CASO	38
3.1 – Introdução	38
3.2 – Arquitetura do Sistema de Telefonia da UFU.....	38
3.2.1 – PABX	40
3.2.2 – Voice Gateway	41
3.2.3 – CDR.....	44
3.2.3.1 – CDR do Gateway de voz.....	45
3.2.3.2 – CDR do PABX.....	47
3.3 – Considerações Finais	48
CAPÍTULO 4 – MODELAGEM E SIMULAÇÃO DE UM GATEWAY DE VOZ	49
4.1 – Introdução	49
4.2 – Modelo de Simulação	49
4.3 – Coleta de Dados	50
4.4 – Modelagem da Variável "Tempo entre Chegada de Chamadas"	52
4.5 – Modelagem da Variável "Duração da Chamada"	54
4.6 – Variáveis de Resposta da Simulação	55
4.7 – Cenários Simulados	56
4.8 – Ferramentas Utilizadas na Modelagem e Simulação	57
4.8.1 – Software Arena	57
4.8.1.1 – Modelagem da "Duração da Chamada" e “Recursos” no Arena	59
4.8.1.2 – Modelagem do "Tempo entre Chegada de Chamadas" no Arena	60
4.8.2 – Input Analyser	62
4.8.3 – Minitab	63
4.8.4 – Microsoft Excel	63
4.9 – Validação do Modelo.....	63
4.9.1 – Validação do Modelo de Simulação Baseada nos Cenários #1 e #2.....	65
4.10 – Considerações Finais	66
CAPÍTULO 5 – ANÁLISE DOS RESULTADOS DA SIMULAÇÃO	68
5.1 – Resultados do Cenário #1	68

5.2 – Resultados do Cenário #2	69
5.3 – Resultados do Cenário #3	70
5.4 – Resultados do Cenário #4	71
5.5 – Resultados do Cenário #5	72
5.6 – Resultados do Cenário #6	73
5.7 – Considerações Finais	74
CAPÍTULO 6 – CONCLUSÕES DA PESQUISA.....	76
6.1 – Principais Resultados	76
6.2 – Limitações da Pesquisa	77
6.3 – Dificuldades Encontradas	77
6.4 – Contribuição para a Literatura	78
6.5 – Trabalhos Futuros	78
REFERÊNCIAS BIBLIOGRÁFICAS	79

LISTA DE TABELAS

Tabela 2.1 - Valores críticos para o teste K-S.	27
Tabela 2.2 - Métodos SIP.	30
Tabela 2.3 - Classes de respostas SIP.....	31
Tabela 3.1 - Campos de um CDR no <i>gateway</i> de voz.....	45
Tabela 3.2 - CDR de uma chamada atendida completada via provedor PSTN.....	46
Tabela 3.3 - CDR de uma chamada atendida encaminhada para o provedor VoIP.	46
Tabela 3.4 - CDR de uma chamada não atendida encaminhada para o provedor PSTN.	47
Tabela 3.5 - Campos de um CDR de chamadas para celular coletado no PABX.	48
Tabela 4.1 - Recursos do <i>gateway</i> utilizados por cada tipo de chamada.....	50
Tabela 4.2 - Estatísticas descritivas para o número de chamadas.	51
Tabela 4.3 - Estatísticas descritivas para a duração das chamadas (em segundos).	51
Tabela 4.4 - Taxa média de chegada de chamadas dos 30 dias com maior volume de chamadas agrupadas em períodos de 15 minutos.....	52
Tabela 4.5 - Resultados do teste A-D para a "Duração da Chamada".....	54
Tabela 4.6 - Resultados do teste χ^2 para a "Duração da Chamada".....	54
Tabela 4.7 - Parâmetros da distribuição LogNormal para a variável "Duração da chamada".	55
Tabela 4.8 - Variáveis de reposta obtidas por simulação.	56
Tabela 4.9 - Tipos de chamadas simuladas e recursos utilizados.....	57
Tabela 4.10 - Resultados simulados e observados do cenário #1.....	65
Tabela 4.11 - Resultados simulados e observados do cenário #2.....	66
Tabela 5.1 - Resultado da simulação do Cenário #1.	68
Tabela 5.2 - Resultado da simulação do Cenário #2.	69
Tabela 5.3 - Resultado da simulação do Cenário #3.	71
Tabela 5.4 - Resultado da simulação do cenário #4.	72
Tabela 5.5 - Resultado da simulação do Cenário #5.	73
Tabela 5.6 - Resultado da simulação do Cenário #6.	74

LISTA DE FIGURAS

Figura 2.1 - Estrutura genérica de uma mensagem SIP.....	30
Figura 2.2 - Exemplo de troca de mensagens entre dois clientes SIP.	31
Figura 3.1 - Visão geral da atual arquitetura do sistema de telefonia da UFU.....	39
Figura 3.2 - Visão geral do sistema de telefonia da UFU com cartões GSM.....	40
Figura 3.3 - Visão frontal do PABX da UFU.....	40
Figura 3.4 - Conexão física do PABX com o <i>gateway</i> de voz utilizando dois <i>modems</i>	42
Figura 3.5 - Visão frontal do <i>modem</i>	42
Figura 3.6 - Visão da parte traseira do <i>modem</i> que é ligado ao servidor Asterisk (cabo UTP com conector RJ-45).....	42
Figura 3.7 - Visão da parte traseira do <i>modem</i> que é ligado ao PABX (cabo coaxial com conector BNC).....	43
Figura 3.8 - Visão frontal do Servidor XVM-VOIP.....	43
Figura 3.9 - Visão traseira do servidor XVM-VOIP, mostrando a ligação do cabo com conector RJ45 vindo do <i>modem</i> para a placa PCI do servidor.	43
Figura 3.10 - Configuração do canal DAHDI no Asterisk.....	44
Figura 3.11 - Configuração da placa E1 PCIe no servidor XVM-VOIP.....	44
Figura 4.1 - Visão geral do modelo de simulação.	49
Figura 4.2 - Taxa média de chegada de chamadas por períodos de 15 minutos.	53
Figura 4.3 - Gráfico de probabilidade (<i>p-p plot</i>) da variável aleatória "Duração da Chamada" para o tipo de chamada II (PSTN).	55
Figura 4.4 - Gráfico de probabilidade (<i>p-p plot</i>) da variável aleatória "Duração da Chamada" para o tipo de chamada III (GSM).....	55
Figura 4.5 - Modelo construído no Arena.	58
Figura 4.6 - Tela de configuração do tempo de processamento e recursos.	60
Figura 4.7 - Tela de configuração do tempo entre chegada de chamadas.	60
Figura 4.8 - Editor gráfico de <i>Schedule</i> do software Arena.	61
Figura 4.9 - Exemplo do resultado de uma simulação apresentado pelo Arena.....	62
Figura 4.10 - Teste de aderência utilizando o Input Analyser.....	63
Figura 4.11 - Visão geral do processo de validação do modelo de simulação.....	64

LISTA DE ABREVIATURAS E SIGLAS

A-D	Teste de aderência Anderson-Darling
ARIMA	Autoregressive Integrated Moving Average
CDF	Cumulative Density Function
CDR	Call Detail Record
CSV	Comma-Separated Values
DDD	Discagem Direta a Distância
DDI	Discagem Direta Internacional
GNA	Gerador de Números Aleatórios
GSM	Global System for Mobile
HTTP	Hypertext Transfer Protocol
HW _{CI}	Half-Width Confidence Interval
IC	Intervalo de Confiança
IFES	Instituição Federal de Ensino Superior
IP	Internet Protocol
ISDN	Integrated Services Digital Network
ITU	International Telecommunication Union
ITU-T	ITU Telecommunication Standardization Sector
K-S	Teste de aderência Kolmogorov-Smirnov
LCG	Linear Congruential Generator
LHS	Latin Hypercube Sampling
NHPP	Non-Homogeneous Poisson Process
ODBC	Open Database Connectivity
PABX	Private Automatic Branch Exchange
PCI	Peripheral Component Interconnect
P-P Plots	Probability-Probability Plots
PSTN	Public Switched Telephone Network
QED	Qualify-Efficiency Driven
QoS	Quality of Service
RFC	Request for Comments
RNP	Rede Nacional de Ensino e Pesquisa
RTCP	Real-time Transport Control Protocol

RTP	Real-time Transport Protocol
SIP	Session Initiation Protocol
SLA	Service Level Agreement
SQL	Structured Query Language
TCP	Transmission Control Protocol
TI	Tecnologia da Informação
UDP	User Datagram Protocol
UFU	Universidade Federal de Uberlândia
URA	Unidade de Resposta Audível
UTP	Unshielded Twisted Pair
VoIP	Voice over IP
χ^2	Teste de aderência Qui-Quadrado

CAPÍTULO 1 – INTRODUÇÃO

1.1 – Contextualização

Embora a comunicação através das redes sociais, e-mail e mensagem instantânea tenha crescido nas últimas décadas, o telefone ainda é essencial em vários ambientes organizacionais.

A integração da telefonia convencional com a rede de dados, utilizando protocolos VoIP (*Voice over Internet Protocol*) [Goode 2002], permite que as chamadas telefônicas sejam completadas a partir do uso da infraestrutura da rede de dados (Internet). A utilização da tecnologia VoIP permite: *i*) a interconexão entre os vários escritórios da organização sem precisar contratar serviços específicos de telecomunicação para esse fim; *ii*) chamadas DDD e internacionais através da Internet; e *iii*) possibilidade de utilizar o mesmo ramal da organização em diferentes locais (mobilidade).

Geralmente, a implementação de VoIP inicia-se mesclando a telefonia convencional com a telefonia VoIP, pois a substituição completa da infraestrutura legada requer alto investimento e não ocorre em curto espaço de tempo. Nas grandes organizações, é comum que o primeiro passo para a adoção de VoIP seja a instalação de um sistema *gateway* de voz entre a rede de telefonia convencional e a rede de telefonia IP para permitir que usuários da telefonia convencional utilizem os serviços oferecidos pela telefonia IP.

Devido ao rápido crescimento da utilização da telefonia IP, e o papel importante dos sistemas *gateways* de voz nessa transição, é fundamental que as organizações planejem a capacidade desses sistemas para garantir a qualidade do serviço de telefonia para o usuário final. Este planejamento de capacidade deve ser realizado proativamente, para que os serviços oferecidos possam cumprir o SLA (*Service Level Agreements*) para a demanda (corrente e futura) de chamadas.

Visando contribuir com as pesquisas nessa área, este trabalho propõe uma abordagem baseada em simulação para o planejamento de capacidade de sistemas *gateway* de voz. Essa abordagem é avaliada realizando a análise e o planejamento da capacidade de um sistema *gateway* de voz real, o qual é usado em uma IFES composta por seis campi. Neste estudo a modelagem e simulação computacional foram usadas para planejar os recursos críticos desse sistema.

1.2 – Relevância do Trabalho

O processo de planejamento de capacidade de recursos de TI (Tecnologia da Informação) visa garantir que exista uma capacidade adequada de recursos, a um custo justificado, e que esta capacidade esteja alinhada com as necessidades atuais e futuras do negócio [Freitas 2010].

Em sistemas de telefonia, as soluções analíticas ainda são frequentemente utilizadas para dimensionar a quantidade de recursos necessária para atender determinado volume de ligações. No entanto, as pesquisas recentes vêm apresentando alternativas aos modelos analíticos, principalmente porque muitas das suposições necessárias para suas aplicações não ocorrem na prática, o que torna os seus resultados questionáveis.

Alguns trabalhos (ex. [Franzese *et al.* 2009] e [Bouzada 2009]) propõem a utilização de modelos de simulação como alternativa aos modelos analíticos para fazer o dimensionamento da capacidade de sistemas de telefonia corporativa (ex. *call centers* e centrais de atendimento). No modelo de simulação é possível incluir detalhes que não seria conveniente incluir nos modelos analíticos, principalmente devido ao aumento na complexidade da solução que estes implicariam. Nos trabalhos que propõem a utilização de simulação para planejar capacidade de recursos em sistemas de telefonia corporativa, não foram encontradas pesquisas específicas que tratam do planejamento de capacidade de recursos de sistemas *gateways* de voz.

Por se tratar de sistemas que têm a possibilidade de entroncamentos com diferentes operadoras de telefonia, possuindo diferentes tipos de recursos necessários para completar chamadas e atender diferentes perfis de chamadas simultaneamente, o nível de complexidade envolvido no planejamento destes sistemas torna difícil o uso de modelos analíticos para o planejamento de capacidade de sistemas *gateway* de voz.

1.3 – Objetivos da Pesquisa

1.3.1 – Objetivo Geral

Sistematizar a modelagem e simulação de sistemas *gateway* de voz, visando ao planejamento de capacidade para demandas atuais e futuras.

1.3.2 – Objetivos Específicos

- Especificar e implementar um modelo estocástico de simulação para um sistema *gateway* de voz real, voltado para o planejamento de sua capacidade.

- Validar o modelo de simulação proposto, por meio de análise estatística de significância dos seus resultados contra valores reais observados em campo.
- Avaliar a robustez do modelo de simulação para sua aplicação em diferentes cenários.
- Avaliar a propriedade e alcance dos resultados obtidos.

1.4 – Desenvolvimento da Pesquisa

Neste trabalho foi realizado um estudo de natureza teórica e experimental, cujos passos estão descritos a seguir.

1.4.1 – Revisão da Literatura

Nesta etapa, foram estudados artigos relacionados com o tema “planejamento de capacidade de sistemas de telefonia corporativa” e correlatos, os quais foram publicados em revistas e conferências especializadas, em nível nacional e internacional.

Também, foram estudados os assuntos modelagem e simulação computacional, bem como técnicas de análise de dados e inferência estatística.

1.4.2 – Definição do Estudo de Caso

Durante esta pesquisa, foi realizado um estudo de caso utilizando o sistema de telefonia da Universidade Federal de Uberlândia (UFU). Além de interconectar as extensões internas locais, este sistema de telefonia se conecta a três provedores de telefonia: cabeada (PSTN), wireless (GSM) e VoIP.

O provedor VoIP é utilizado para fazer ligações DDD com tarifas reduzidas. As ligações que não podem ser completadas via VoIP são encaminhadas para a operadora de telefonia PSTN (*Public Switched Telephone Network*).

Atualmente, o *gateway* de voz modelado como estudo de caso é o sistema responsável por interconectar o PABX da UFU com o provedor VoIP. A sua principal função é traduzir os protocolos ISDN (telefonia convencional) para os protocolos utilizados na telefonia IP. Atualmente esse sistema atende 777 usuários, processando em média 1034 ligações por dia.

Dada a importância do *gateway* de voz no sistema de telefonia da UFU, este foi selecionado para o estudo de caso do presente trabalho.

1.4.3 – Coleta de Dados

Nesta etapa, os registros detalhado das chamadas (CDR) do *gateway* de voz e do PABX da UFU foram coletados. Foram utilizados os dados de Janeiro/2012 até Dezembro/2012 para servir de entrada no modelo de simulação.

1.4.4 – Análise dos Dados de Entrada

Nesta etapa, foram aplicados testes estatísticos de aderência aos dados coletados, a fim de verificar o melhor modelo de distribuição de probabilidades para representar os dados em questão.

Os seguintes testes de aderência foram utilizados:

- Kolmogorov-Smirnov (K-S)
- Anderson-Darling (A-D)
- Qui-Quadrado (χ^2)

Além dos testes de aderência, também foi utilizada a análise gráfica (*p-p plots*) durante a etapa de identificação das distribuições de probabilidades usadas no modelo de simulação.

1.4.5 – Criação e Validação do Modelo de Simulação

Nesta etapa, criou-se o modelo de simulação para representar o sistema de telefonia investigado. Para validar o modelo, contou-se com o apoio dos técnicos especialistas que são responsáveis pela operação do sistema de telefonia modelado. Além disso, validou-se quantitativamente o modelo utilizando os dados reais de ligações obtidos dos sistemas *gateway* e PABX atualmente em funcionamento.

1.4.6 – Execução das Simulações

Nesta etapa, as simulações foram executadas para verificar o comportamento do sistema com variadas configurações de recursos. Diferentes fases de calibração do modelo de simulação foram necessárias para ajustar os parâmetros do modelo a fim de torná-lo compatível com os dados reais observados.

1.4.7 – Análise dos Resultados

Nesta etapa, os resultados das simulações foram analisados com foco em diferentes variáveis de interesse, a fim de entender e avaliar o comportamento do sistema *gateway*

simulado no estudo de caso, com vistas para diferentes cenários de planejamento de capacidade.

1.5 – Estrutura do Trabalho

Os demais capítulos desta dissertação estão organizados como segue.

O Capítulo 2 apresenta os conceitos teóricos de modelagem e simulação computacional. Também são apresentados trabalhos da literatura que abordam o assunto de planejamento de capacidade de sistemas de telefonia corporativos.

O Capítulo 3 descreve o sistema de telefonia utilizado no estudo de caso, enfatizando o PABX e as suas conexões com as operadoras de telefonia. Além disso, são apresentados diversos detalhes do sistema *gateway* estudado.

O Capítulo 4 apresenta a abordagem proposta para modelar e simular um *gateway* de voz, detalhando os métodos e materiais adotados.

O Capítulo 5 apresenta os resultados dos diversos cenários simulados no estudo de caso, bem como a análise do planejamento de capacidade do *gateway* de voz avaliado.

Por fim, o Capítulo 6 apresenta as conclusões da pesquisa, as dificuldades encontradas, as suas limitações, as contribuições para a literatura e as sugestões de trabalhos futuros.

CAPÍTULO 2 – REVISÃO DA LITERATURA

2.1 – Introdução

Este capítulo está dividido em três partes. Na primeira parte (Seção 2.2), são apresentados os conceitos teóricos sobre modelagem, simulação e estatística que foram utilizados durante a pesquisa. Na segunda parte (Seção 2.3), são apresentados os protocolos SIP e RTP, usados em sistemas *gateway* de voz baseados em VoIP. Na terceira parte (Seção 2.4), são apresentados os trabalhos encontrados na literatura relacionados com o planejamento de capacidade de recursos de sistemas de telefonia corporativa.

2.2 – Modelagem e Simulação

Em [Kelton *et al.* 2007], os autores definem simulação como uma ampla coleção de métodos e aplicações que pode ser utilizada para imitar o comportamento de sistemas reais, geralmente usando um computador com software apropriado. Do ponto de vista prático, simulação é o processo de concepção e criação de um modelo computacional de um sistema real, ou de um sistema proposto, com o objetivo de realizar experimentos numéricos para entender melhor o comportamento do sistema. Embora seja uma técnica que pode ser utilizada para estudar sistemas simples, o verdadeiro poder da simulação é plenamente alcançado quando ela é utilizada para estudar sistemas complexos [Kelton *et al.* 2007] [Sanchez 2007].

A simulação é utilizada em diversas áreas para analisar sistemas que são complicados de serem estudados via métodos analíticos, tais como cálculo, probabilidade e estatística, ou teoria das filas [Goldsman 2007]. Na maioria das vezes, a simulação é utilizada para medir o desempenho, melhorar a operação ou projetar um sistema caso ele ainda não exista [Kelton *et al.* 2007] [Law e Kelton 1991].

A seguir é apresentada uma lista de problemas em que a simulação pode ser uma ferramenta útil para ajudar nas suas resoluções [Law e Kelton 1991]:

- Projetar e analisar sistemas de produção;
- Avaliar requisitos de hardware e software para um computador;
- Avaliar novas táticas militares;
- Projetar sistemas de comunicação e seus protocolos;
- Projetar e operar instalações de transporte, como rodovias, aeroportos, metrô ou portos;

- Avaliar projetos para organização de serviços, tais como hospitais, correios ou restaurantes;
- Analisar sistemas financeiros e econômicos.

Os modelos de simulação podem ser classificados em discretos ou contínuos. Nos modelos de simulação contínua, a modelagem do sistema ao longo do tempo é realizada por uma representação em que as variáveis de estado alteram-se de forma contínua em relação ao tempo. Geralmente, os modelos de simulação contínuos envolvem a utilização de equações diferenciais. Já os modelos de simulação discreta referem-se à modelagem de um sistema ao longo do tempo por uma representação em que as variáveis de estado do sistema são alteradas em pontos separados do tempo.

Além da classificação em discretos ou contínuos, os modelos de simulação também podem ser classificados em determinísticos ou estocásticos. Os modelos determinísticos não contêm componente probabilístico (aleatório). Já nos modelos de simulação estocásticos existe pelo menos um componente aleatório como entrada. Portanto, os modelos estocásticos fornecem resultados que são aleatórios. Assim, esses resultados precisam ser tratados apenas como uma estimativa dos resultados reais.

Neste trabalho foi utilizada a simulação discreta e estocástica. A seguir serão apresentados mais detalhes sobre esse tipo de abordagem.

2.2.1 – Simulação Discreta de Eventos

A simulação discreta de eventos é caracterizada pela alteração do estado do sistema em pontos separados do tempo. Em termos matemáticos, pode-se dizer que o sistema só pode sofrer alteração em um número de pontos enumeráveis no tempo. Esses pontos no tempo são aqueles em que um evento ocorre.

O termo evento é definido como uma ocorrência instantânea que pode alterar o estado do sistema. Embora este tipo de simulação possa ser feita manualmente, a quantidade de dados para armazenar e manipular, para a maioria dos sistemas reais, exige que a simulação seja feita por computador [Law e Kelton 1991].

Devido a sua natureza dinâmica, os modelos de simulação discreta de eventos precisam manter o valor corrente do tempo de simulação e de um mecanismo para avançar o tempo de simulação de um valor para outro. A próxima seção (2.2.1.1) descreve dois mecanismos de avanço no tempo: *next-event time-advance* e *fixed-increment time-advance*.

2.2.1.1 – Mecanismos de Avanço do Tempo

Na simulação discreta de eventos é necessário guardar o valor do tempo enquanto a simulação ocorre. Também é necessário um mecanismo que avança no tempo de um valor para outro. A variável que guarda o valor atual do tempo da simulação é chamada “relógio da simulação” [Law e Kelton 1991].

Existem dois tipos de avanço do relógio de simulação: *next-event time-advance* e *fixed-increment time-advance*. No primeiro, o relógio da simulação é iniciado com zero e os horários de ocorrência dos eventos futuros são determinados. O relógio da simulação é então incrementado para o horário da ocorrência do primeiro dos eventos futuros. Nesse momento, o estado do sistema é atualizado e os horários dos eventos futuros são determinados.

Uma vez que todas as mudanças de estado na simulação discreta de eventos acontecem apenas quando os eventos ocorrem, os períodos de inatividade são ignorados fazendo o relógio pular de *event-time* em *event-time*.

No tipo de avanço *fixed-increment time advance*, o relógio da simulação é sempre incrementado de um valor Δt escolhido. Após cada avanço do relógio, é feita a verificação para descobrir se algum evento ocorreu durante esse intervalo (*time* - Δt). Caso tenham ocorrido eventos, eles serão executados no final do intervalo e as variáveis que mantêm o estado do sistema são atualizadas [Law e Kelton 1991].

Em [Law e Kelton 1991], os autores destacam duas desvantagens desse tipo de abordagem:

- Erros pelo fato de o evento ser executado apenas no final do intervalo;
- Escolha de qual evento deverá ser executado primeiro quando os eventos não são simultâneos no sistema real.

Essas desvantagens podem ser minimizadas com a escolha do menor valor Δt . No entanto, essa operação provoca aumento no tempo de processamento da lista de eventos. Portanto, o tipo de avanço *fixed-increment* geralmente não é utilizado quando o tempo entre dois eventos sucessivos varia muito.

Neste trabalho foi utilizado o mecanismo de avanço no tempo *next-event time advance*.

2.2.1.2 – Números Aleatórios e Variáveis Aleatórias

Esta seção descreve o que é um GNA (Gerador de números aleatórios) e apresenta dois exemplos de geração de variáveis aleatórias para a distribuição exponencial e para a distribuição uniforme.

O gerador de números aleatórios é um elemento importante para qualquer simulação estocástica. O único propósito do GNA é produzir um conjunto de números que são observações de uma distribuição uniforme contínua entre 0 e 1 e que são independentes uns dos outros [Kelton *et al.* 2007].

Uma forma bastante usada para gerar números aleatórios é através do Método Congruente Linear Multiplicativo (MCLM) [Freitas Filho 2008], que é derivado do Método Congruente Misto. A Equação 1 mostra uma função geradora do tipo MLCM:

$$x_i = ax_{i-1} \bmod m, \quad (1)$$

onde a e m são chamados de multiplicador e módulo respectivamente e precisam ser escolhidos com cuidado, baseado tanto em motivos teóricos quanto em empíricos, para produzir um bom conjunto de números aleatórios.

O processo de geração de variáveis aleatórias de qualquer distribuição começa com a geração de números aleatórios da distribuição uniforme contínua entre 0 e 1 [Law e Kelton 1991]. Nos modelos de simulação estocásticos, as variáveis aleatórias são utilizadas para diferentes propósitos. Elas geralmente servem como modelo de entrada de quantidades como duração incerta de tempos (ex. serviços ou tempos entre chegadas) ou número de clientes de um grupo de chegadas [Kelton *et al.* 2007].

Dado um conjunto de números aleatórios, em princípio, podemos usar a função inversa da CDF (*Cumulative Density Function*) para gerar variáveis aleatórias de qualquer distribuição. No entanto, não é possível usar a função inversa da CDF quando não há equação, em forma fechada, para fazer o cálculo da CDF. A distribuição normal é um exemplo onde isso ocorre. Neste caso, as variáveis aleatórias podem ser geradas usando *i*) convolução, *ii*) composição ou *iii*) transformações baseadas no relacionamento entre distribuições particulares (ex. tabela da normal padrão) [Sanchez 2007].

Existem vários algoritmos que podem ser utilizados para gerar variáveis aleatórias de uma dada distribuição, e vários fatores precisam ser considerados na escolha do algoritmo que será utilizado em um estudo particular de simulação. Como alguns fatores são conflitantes entre si, o analista deve escolher o que for mais apropriado para determinada situação. Em [Law e Kelton 1991], os autores citam que os fatores exatidão, eficiência, complexidade e robustez devem ser considerados no processo de escolha do algoritmo.

A seguir serão apresentados dois exemplos de geração de variáveis aleatórias. O primeiro para a distribuição exponencial e o segundo para a distribuição uniforme.

A função de densidade acumulada, $F(x)$, da distribuição exponencial é dada por:

$$F(x) = 1 - e^{\frac{-x}{\beta}}. \quad (2)$$

Para gerar uma variável aleatória X , que segue a distribuição exponencial, inicia-se gerando um número aleatório U da distribuição uniforme contínua entre 0 e 1. Em seguida faz-se U igual à CDF e resolve-se X em termos do valor de U (agora conhecido). Essa solução de X em termos de U é chamada de CDF inversa de U , uma vez que essa transformação desfaz para U o que F faz para X [Kelton *et al.* 2007].

Para gerar uma variável aleatória para a distribuição uniforme dentro do intervalo $[a; b]$, pode-se gerar um número aleatório U da distribuição uniforme contínua entre 0 e 1, e realizar o seguinte cálculo [Law e Kelton 1991]:

$$X = a + U \times (b - a). \quad (3)$$

Neste trabalho as variáveis aleatórias foram utilizadas para representar o tempo entre as chegadas de chamadas telefônicas e a duração das chamadas.

2.2.2 – Linguagens de Simulação

Para desenvolver a maioria dos modelos de simulação discreta, a linguagem de simulação deve fornecer algumas funcionalidades básicas. Em [Law e Kelton 1991], os autores destacam que as seguintes funcionalidades são desejáveis para criar projetos de simulação:

- Gerador de valores aleatórios de uma distribuição específica;
- Avançar o tempo de simulação;
- Determinar o próximo evento da lista de eventos e passar o controle para o bloco de código apropriado;
- Manipular listas;
- Coletar e analisar dados;
- Gerar relatórios com os resultados;
- Detectar condições de erro;

A maioria das linguagens de simulação oferece esses recursos sem a necessidade de programação adicional, tornando o código mais simples e o desenvolvimento mais rápido. Como o código fica menor, a probabilidade de erros também é menor. No entanto, a decisão de usar uma linguagem de propósito geral ou uma linguagem de simulação para desenvolver o estudo deve levar em conta outras características tais como *i*) conhecimento dos analistas em

determinadas linguagens; *ii*) rapidez na execução da simulação; e *iii*) custo do software [Law e Kelton 1991].

Uma vez decidido que será utilizada uma linguagem de simulação, o analista deve verificar quais recursos as linguagens oferecem para definir sua escolha. Em [Law e Kelton 1991], os autores listam um conjunto de funcionalidades, divididas em cinco categorias, as quais podem ser utilizadas para ajudar na escolha da linguagem adequada:

- Funcionalidades gerais, como capacidade de gerar números aleatórios para diferentes distribuições de probabilidade e ter um único comando para executar as replicações dos modelos;
- Recursos de animação;
- Recursos de análise estatística dos resultados;
- Suporte ao usuário;
- Relatórios dos resultados.

2.2.3 – Testes de Aderência

Nos modelos de simulação estocásticos, é necessário representar os dados de entrada do modelo por meio de distribuições de probabilidade. Para isso, faz-se necessário identificar, dentre as diferentes distribuições teóricas consideradas, qual melhor adere aos dados analisados. Para tanto, utiliza-se os chamados testes de aderência (*Goodness of fit test*) [Barbetta *et al.* 2010].

O objetivo de um teste de aderência é verificar se os dados de uma amostra comportam-se de acordo com uma distribuição teórica de probabilidade [Barbetta *et al.* 2010]. Neste tipo de teste, admite-se, por hipótese, que a distribuição da variável de interesse na população seja descrita por determinado modelo de distribuição de probabilidade, ou seja, verifica-se a boa ou má aderência de distribuições de probabilidades teóricas aos dados da amostra. Caso obtenha-se uma boa aderência e a amostra for razoavelmente grande, pode-se, em princípio, admitir que o modelo forneça uma boa representação da distribuição populacional. Inversamente, a rejeição da hipótese em um dado nível de significância indica que o modelo testado é inadequado para representar a distribuição da população [Costa Neto 1977].

É comum encontrar na literatura que os testes de aderência testam as seguintes hipóteses:

H_0 : Os dados seguem uma distribuição específica.

H_1 : Os dados não seguem uma distribuição específica.

Neste trabalho, foram utilizados três testes de aderência, a saber: Kolmogorov-Smirnov (K-S) [Massey 1951], Anderson-Darling (A-D) [Anderson e Darling 1954] e Qui-Quadrado (χ^2) [Boero *et al.* 2005]. Também, de forma complementar aos testes numéricos, foi realizada a análise gráfica da aderência das distribuições por meio de gráficos de probabilidade (*p-p plots*). A seguir serão abordados os três testes de aderência que foram utilizados durante a pesquisa.

2.2.3.1 – Kolmogorov-Smirnov

Kolmogorov e Smirnov desenvolveram um teste de aderência, não paramétrico, em que a variável de teste é a maior diferença observada entre a função de distribuição acumulada do modelo e a amostra [Costa Neto 1977]. Além de não ser necessário o agrupamento dos dados, como é realizado no teste de aderência χ^2 , o teste K-S tende a ser mais poderoso do que o χ^2 [Law e Kelton 1991].

A função de distribuição acumulada do modelo testado fornece as probabilidades acumuladas em cada ponto, ou seja,

$$F(x) = P(X \leq x). \quad (4)$$

A função de distribuição acumulada da amostra corresponderá ao gráfico das frequências relativas acumuladas. Designando essa segunda função por $G(x)$, o teste consta do cálculo do valor d , realizado por meio da Equação 5,

$$d = \max |F(x) - G(x)|, \quad (5)$$

e da comparação desse valor com um valor crítico tabelado em função do nível de significância (α) e do número de elementos da amostra (n). Se d for maior que o valor crítico, rejeita-se H_0 [Costa Neto 1977].

A Tabela 2.1 apresenta os valores críticos para o teste K-S para níveis de significância $\alpha=1\%$ e $\alpha=5\%$. Para $n > 50$, o cálculo dos valores críticos para $\alpha=5\%$ e $\alpha=1\%$ é realizado por, respectivamente, $1,36/\sqrt{n}$ e $1,63/\sqrt{n}$ [Costa Neto 1977] [Barbetta *et al.* 2010].

2.2.3.2 – Anderson-Darling

O teste Anderson-Darling (A-D) é usado para verificar se uma amostra de dados vem de uma população com uma distribuição de probabilidade conhecida. Apesar de o teste também ser feito entre a distribuição teórica acumulada e a distribuição empírica, o teste A-D dá maior peso para a calda do que o K-S [NIST/SEMATECH 2014].

Tabela 2.1 - Valores críticos para o teste K-S.

N	$\alpha=5\%$	$\alpha=1\%$	N	$\alpha=5\%$	$\alpha=1\%$
1	0,975	0,995	14	0,349	0,418
2	0,842	0,929	15	0,338	0,404
3	0,708	0,829	16	0,327	0,392
4	0,624	0,734	17	0,318	0,381
5	0,563	0,669	18	0,309	0,371
6	0,519	0,617	19	0,301	0,361
7	0,483	0,576	20	0,294	0,352
8	0,454	0,542	25	0,264	0,317
9	0,430	0,513	30	0,242	0,290
10	0,409	0,490	35	0,224	0,269
11	0,391	0,468	40	0,210	0,252
12	0,375	0,449	45	0,198	0,238
13	0,361	0,432	50	0,188	0,227

Os valores críticos do teste K-S são independentes de distribuição. Já no A-D, cada distribuição tem um conjunto de valores críticos. A vantagem de ser dependente de distribuição é que o A-D torna-se mais sensível que o K-S. A desvantagem é que os valores críticos precisam ser calculados para cada distribuição [NIST/SEMATECH 2014].

Assim como no teste K-S, o teste A-D testa as seguintes hipóteses:

H_0 : Os dados seguem uma distribuição específica.

H_1 : Os dados não seguem uma distribuição específica.

A estatística do teste, A^2 , é definida por (6).

$$A^2 = -N - S, \quad (6)$$

onde S é calculado através da Equação 7,

$$S = \sum_{i=1}^n \frac{2i-1}{n} \times [\ln F(Y_i) + \ln(1 - F(Y_N + 1 - i))], \quad (7)$$

e F é a função de densidade acumulada.

As tabelas com valores críticos do teste A-D, para várias distribuições, podem ser encontradas em [Stephens 1974], [Stephens 1976], [Stephens 1977a], [Stephens 1977b], [Stephens 1979].

2.2.3.3 – Qui-Quadrado

O teste de aderência Qui-Quadrado (χ^2) pode ser aplicado quando temos interesse em estudar dados distribuídos em categorias e deseja-se verificar se as frequências observadas em K diferentes categorias ($O_i, i=1, 2, \dots, K$) são significativamente distintas de um conjunto de K frequências esperadas ($E_i, i=1, 2, \dots, K$) [Barbetta *et al.* 2010]. As hipóteses são:

H_0 : $O_i = E_i$ para todo $i = 1, 2, \dots, K$.

H_1 : $O_i \neq E_i$ para algum $i = 1, 2, \dots, K$.

A estatística χ^2 do teste Qui-Quadrado é uma espécie de medida de distância entre as frequências observadas e as frequências esperadas. Para calcular χ^2 utiliza-se a Equação 8.

$$\chi^2 = \sum_{i=1}^K \left[\frac{(O_i - E_i)^2}{E_i} \right]. \quad (8)$$

Caso o valor de χ^2 seja pequeno, significa que as frequências observadas são próximas das frequências esperadas. Neste caso, considera-se H_0 verdadeira, pois as variações encontradas seriam apenas casuais. No caso em que o valor de χ^2 é grande, significa que as diferenças entre frequências observadas e esperadas podem ser grandes. Assim, nesse caso, é pouco provável que as variações tenham sido casuais.

2.2.4 – Gráficos de Probabilidade

Quando se trabalha com amostras de tamanho grande, não é incomum que os testes de aderência rejeitem a hipótese nula [Law e Kelton 1991]. Portanto, neste trabalho também foi utilizada a análise gráfica (*p-p plots*) para verificar a aderência das amostras de chamadas telefônicas a alguma distribuição teórica de probabilidade.

O gráfico de probabilidade [Chambers *et. al.* 1983] é uma técnica gráfica que pode ser utilizada para verificar se um conjunto de dados segue uma distribuição de probabilidade teórica.

2.2.5 – Intervalo de Confiança para a Média

Como os resultados da simulação são estocásticos, é necessário calcular o intervalo de confiança para as medidas (ex. média) dos indicadores de desempenho de interesse.

Um intervalo de confiança (IC) para um indicador de desempenho de interesse compreende um intervalo numérico que possui uma probabilidade igual a $100 \times (1 - \alpha)\%$ de incluir o verdadeiro valor da medida sob análise [Freitas Filho 2008], onde α é o erro admitido ao se concluir sobre a presença do verdadeiro valor da variável no intervalo calculado.

Para obter o intervalo de confiança é necessário calcular a distância, conhecida como semi-intervalo de confiança, simbolizada pela letra h , que determina os limites inferiores e superiores do intervalo. Logo, um intervalo de confiança para uma média estimada \bar{X} é dado por (9),

$$IC = [\bar{X} - h; \bar{X} + h]. \quad (9)$$

Para amostras de tamanho $n \leq 30$, o semi-intervalo h é calculado usando a Equação 10,

$$h = t_{n-1, 1-\alpha/2} \times \frac{S}{\sqrt{n}}, \quad (10)$$

onde $t_{n-1, 1-\alpha/2}$ é o valor tabulado de t (distribuição t de Student) e S é o desvio padrão da amostra.

Para amostras maiores, com $n > 30$, o semi-intervalo é calculado por (11),

$$h = z \times \frac{S}{\sqrt{n}}, \quad (11)$$

onde z é o valor tabelado da distribuição normal padrão e S é o desvio padrão da amostra.

Os três fatores que influenciam a largura do intervalo de confiança são:

- O número de elementos da amostra;
- O nível de confiança;
- A variância associada à medida de desempenho sob análise e capturada pela amostra.

2.3 – Protocolos VoIP

No sistema de telefonia VoIP da UFU, são utilizados dois protocolos principais: o SIP (*Session Initiation Protocol*) para fazer a sinalização, e o RTP (*Real-time Transport Protocol*) para o transporte do tráfego de voz. A seguir esses dois protocolos serão detalhados.

2.3.1 – SIP

O SIP [Rosenberg *et al.* 2002] é um protocolo *text-based* da camada de aplicação da arquitetura TCP/IP. Ele permite estabelecer, modificar e terminar sessões multimídias de tempo real entre os participantes. As sessões podem ser *unicast* (ponto a ponto) ou *multicast* (conferência).

O protocolo SIP é apenas um componente de um conjunto de protocolos necessários para suportar a troca de dados multimídias na Internet que permite que uma entidade faça uma chamada para outra entidade e ambas negociem os parâmetros da sessão multimídia. Ele suporta 5 facetas do processo de estabelecimento e término das comunicações multimídias [Rosenberg *et al.* 2002]:

- Localização de usuário;
- Disponibilidade do usuário;
- Capacidade do usuário;
- Configuração da sessão;
- Gerenciamento da sessão.

O protocolo SIP é baseado no modelo de transação requisição/resposta semelhante ao do protocolo HTTP. A entidade que faz a requisição é chamada de cliente e aquela que recebe a requisição é o servidor. Cada transação consiste em uma requisição realizada pelo cliente que invoca um método (ou função) particular no servidor e pelo menos uma resposta do servidor para o cliente.

As mensagens de requisição e resposta do protocolo SIP consistem em uma linha inicial (*start-line*), um ou mais campos de cabeçalhos, uma linha em branco indicando o final do cabeçalho e o corpo da mensagem (opcional). A Figura 2.1, retirada da RFC 3261, mostra a estrutura genérica de uma mensagem SIP.

generic-message	=	start-line
		* message-header
		CRLF
		[message-body]
Start-line	=	Request-Line / Status-Line

Figura 2.1 - Estrutura genérica de uma mensagem SIP.

O que diferencia uma mensagem de requisição de uma mensagem de resposta é a *start-line*. Uma requisição contém o nome do método, uma *Request-URI* (indicando o usuário ou serviço pelo qual a requisição está sendo endereçada) e a versão do protocolo. Enquanto que uma resposta contém a versão do protocolo seguida por um número que indica o *Status-Code* e o texto relacionado com o status.

A Tabela 2.2 mostra os 6 métodos definidos na RFC 3261. As extensões do protocolo SIP, documentadas em outras RFCs, podem definir métodos adicionais.

Tabela 2.2 - Métodos SIP.

Método	Função
REGISTER	Registrar informações de contato
INVITE	Estabelecer a sessão
ACK	Confirmar o método INVITE
CANCEL	Cancelar a sessão ainda não respondida
BYE	Finalizar a sessão
OPTIONS	Informar a capacidade e disponibilidade dos participantes da sessão

O *Status-Code* de uma mensagem de resposta é um código definido por um número inteiro de 3 dígitos que indica o resultado de uma tentativa para entender e satisfazer uma requisição. O primeiro dígito do *Status-Code* define a classe de resposta. Os outros dois dígitos não possuem nenhuma regra de categorização. A Tabela 2.3 apresenta os seis valores para o primeiro dígito.

A Figura 2.2 ilustra a troca de mensagens de requisições e respostas entre dois clientes SIP. Inicialmente, Alice envia uma mensagem com o método INVITE para iniciar uma sessão com Bob. Enquanto o telefone de Bob toca, duas respostas provisórias (*Status-Code* 100 e 180) são enviadas para Alice. Ao atender a chamada, Bob envia uma mensagem de confirmação (*Status-Code* 200) para Alice. A seguir, Alice envia uma mensagem com o método ACK para Bob para confirmar o estabelecimento da sessão. A partir desse momento, com a sessão estabelecida, existe a troca de tráfego de voz entre Alice e Bob. Para finalizar a chamada, Bob envia uma mensagem com o método BYE para Alice. Ao receber o método BYE, Alice envia uma resposta de confirmação do fim da sessão (*Status-Code* 200).

Tabela 2.3 - Classes de respostas SIP.

Status-Code	Classe	Descrição
1xx	<i>Provisional</i> (Provisória)	A requisição foi recebida; continuando o processamento da requisição.
2xx	Sucesso	A ação foi recebida com sucesso, entendida e aceita.
3xx	Redirecionamento	É necessário tomar mais ações para completar a requisição
4xx	Erro de cliente	A requisição contém erro de sintaxe ou não pode ser completadas neste servidor.
5xx	Erro de servidor	O servidor falhou no atendimento de uma requisição aparentemente válida
6xx	Falha geral	A requisição não pode ser atendida em qualquer servidor.

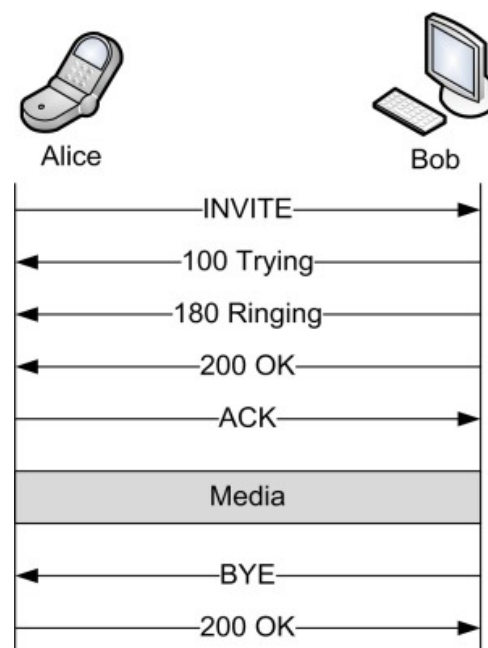


Figura 2.2 - Exemplo de troca de mensagens entre dois clientes SIP.

2.3.2 – RTP

O protocolo SIP é responsável por iniciar, gerenciar e finalizar sessões multimídias. No entanto, ele não oferece suporte para o tráfego de dados. Por isso, pode-se usar o protocolo RTP (*Real-time Transport Protocol*) [Schulzrinne *et al.* 2003] para trafegar os dados multimídia entre os participantes da sessão.

O protocolo RTP, definido pela RFC 3550, oferece o serviço de entrega de dados fim-a-fim em tempo real. Ele permite a transmissão de dados referentes a diferentes tipos de mídias, por exemplo áudio e vídeo, de forma simultânea. Embora possa ser utilizado junto com o protocolo TCP, geralmente é utilizado com o protocolo UDP.

O RTP é responsável por fazer a fragmentação do fluxo de dados multimídia e adicionar a cada fragmento informação de sequência e de tempo de entrega. No entanto, ele não faz reserva de recursos e não garante o QoS (*Quality of Service*) para serviços de tempo real. Ou seja, os pacotes RTP são processados de maneira idêntica aos pacotes de outros protocolos da rede TCP/IP.

Para auxiliar na melhora de qualidade do serviço, o protocolo RTP trabalha em conjunto com o RTCP (*Real-Time Transport Control Protocol*). O RTCP monitora a qualidade do serviço e faz a transmissão periódica de pacotes de controle para todos os participantes da sessão.

2.4 – Trabalhos Correlatos

Após extensa revisão da literatura, não foram encontrados trabalhos que tratam especificamente do planejamento de capacidade de *gateways* de voz, tal como proposto nesta pesquisa. Os trabalhos encontrados que mais se relacionam com esse tema investigam a utilização de técnicas de predição e simulação com o objetivo de modelar e dimensionar a capacidade de *call centers*. Portanto, nesta seção serão apresentados esses trabalhos correlatos.

Em [Mehrotra e Fama 2003], os autores mostram como a simulação computacional pode ser utilizada como ferramenta de apoio para tomada de decisão em *call centers*. Os autores fornecem uma visão geral sobre os principais desafios para criar modelos de simulação de *call centers*. Além disso, é apresentado um exemplo prático de utilização de simulação em *call centers*. Foram simulados vários cenários alterando o número de agentes ativos que poderiam atender ligações receptivas (quando os receptivos estivessem ocupados) e alterando a quantidade de tempo que o cliente espera na fila até fazer o roteamento da chamada para um agente do grupo ativo. O objetivo era achar uma situação que fornecia um

nível de serviço de atendimento de 80% das chamadas entrantes no tempo de 60 segundos com a menor taxa de abandono de chamadas. O desvio padrão, abaixo de 2,5% da média do SLA, foi utilizado para definir o número de replicações de cada cenário. Os autores concluíram que a utilização da simulação foi efetiva para produzir estatísticas detalhadas e variar diferentes parâmetros para examinar os cenários do trabalho, bem como outros que possam ser de interesse.

Em [Green *et al.* 2003], os autores avaliaram a abordagem SIPP (*Stationary Independent Period by Period*) que é amplamente utilizada para estimar o número mínimo de agentes necessários para atender determinada demanda de chamadas em *call centers*. Os autores mostraram que esta abordagem geralmente sugere um número de agentes que é muito baixo para atingir os níveis de serviço durante os períodos com maior volume de ligações. Isso ocorre principalmente porque o SIPP usa a taxa média de chegada por período, assumindo de maneira incorreta que a taxa de chegada é constante durante o período.

Em [Avramidis *et al.* 2004], os autores desenvolveram modelos estatísticos que capturaram duas propriedades observadas em estudos empíricos de processos de chegada de chamadas em *call centers*: chegada de chamadas com intensidade variada ao longo do dia e a correlação diferente de zero entre as quantidades de chamadas em diferentes períodos no mesmo dia. Os autores compararam os novos modelos com os modelos NHPP (*Nonhomogeneous Poisson Process*) e *Doubly Stochastic* (Processo de Poisson com taxa de chegada aleatória). Para cada novo modelo, os autores caracterizaram a distribuição conjunta (*joint distribution*) do vetor de quantidade de chamadas para certificar como os novos modelos são mais flexíveis que os modelos NHPP e *Doubly Stochastic*. Além disso, com os dados de chegadas de um *call center* real foi realizado um estudo para *i*) verificar as características essenciais do processo de chegada de chamadas; *ii*) fazer testes de aderência dos modelos estimados; e *iii*) medir a sensibilidade de várias medidas de desempenho dos modelos usando um modelo de simulação. Eles concluíram que, comparados com modelos tradicionais NHPP (*Non-homogeneous Poisson Process*) e *Doubly Stochastic*, os novos modelos obtiveram melhores precisões nas previsões. No estudo de caso, os novos modelos foram superiores aos tradicionais de duas formas: *i*) a matriz de covariância do vetor de quantidade de chegadas foi melhor aproximada pelos novos modelos; e *ii*) as previsões fornecidas pelos novos modelos foram mais precisas que aquelas fornecidas pelos modelos tradicionais. Além disso, os novos modelos podem ser facilmente adaptados para outras classes de aplicações (ex. e-mail e chat).

Em [Takakuwa e Okada 2005], os autores criaram um modelo de simulação para um *call center* de uma companhia de gás com o objetivo de adequar os horários de intervalo dos atendentes para atender as ligações dentro de um acordo de nível de serviço pré-estabelecido com o menor custo. Foi proposto um procedimento que utilizou programação linear inteira e simulação para descobrir a melhor configuração de horários dos funcionários. A solução inicial foi obtida resolvendo o problema de programação inteira (determinar o número ideal de cada tipo de agente necessário para atender as chamadas dos clientes), e o nível de serviço foi calculado utilizando a simulação. Os autores concluíram que o procedimento proposto foi eficaz em situações reais.

Em [Taylor 2008], os autores avaliaram cinco métodos de predição aplicados a séries univariadas de chegada de chamadas em *call centers* reais. Os resultados obtidos mostraram que os modelos ARIMA sazonais e uma extensão do modelo *Holt-Winters* [Taylor 2003] foram os mais indicados para previsões de até dois ou três dias à frente. Para previsões além de três dias, a média simples foi superior aos demais métodos de previsão.

Em [Robbins e Harrison 2008], os autores apresentam um algoritmo de dois estágios para fazer agendamento de recursos em *call centers* com níveis de serviços rigorosos e taxas de chegada incertas. No primeiro estágio, há a construção heurística de um agendamento, usando um algoritmo guloso, que satisfaz o acordo de nível de serviço. No segundo estágio, a simulação é utilizada para avaliar o agendamento. Em 5 minutos de execução, ou menos, esses dois estágios podem criar um agendamento de recursos com resultados próximos aos agendamentos criados por métodos que necessitam de mais capacidade de processamento computacional.

Em [Franzese *et al.* 2009], os autores apresentam dois modelos analíticos, Erlang B e Erlang C, e modelos de simulação para dimensionar a capacidade de *call centers*. Para comparar os modelos foi realizado um estudo de caso usando um *call center* real de uma companhia de TI (Tecnologia da Informação). O modelo de simulação foi implementado utilizando o software Arena [Arena 2014] e os modelos analíticos foram desenvolvidos no Microsoft Excel. Os autores concluíram que o modelo de simulação ofereceu resultados mais precisos do que os modelos analíticos.

Em [Bouzada 2009], os autores comparam métodos analíticos, baseados em teoria das filas, com métodos experimentais (simulação), quando aplicados no dimensionamento de *call centers* com operações complexas. Foi realizado um estudo de caso com os dados de um *call center* brasileiro. Os autores concluíram que a simulação ofereceu as seguintes vantagens em relação ao método analítico adotado: *i)* permitiu incluir maior número de detalhes das

operações; *ii*) permitiu utilizar diferentes distribuições de probabilidade; *iii*) permitiu o melhor entendimento da operação através da utilização de animações gráficas. Outra conclusão é que o nível de serviço computado pela fórmula Erlang (método analítico), foi subestimado, principalmente porque ele não considera abandono de chamadas.

Em [Aldor-Noiman *et al.* 2009], os autores apresentam um modelo para prever a carga de trabalho e necessidade de funcionários em *call centers* utilizando uma abordagem mista do processo de Poisson [Jongbloed e Koole 2001]. Para avaliar o modelo foram utilizados os dados de um *call center* israelense. Baseado nos critérios para determinar desempenho sobre o regime QED (*Qualify-Efficiency Driven*), os autores concluíram que o modelo proposto atingiu níveis de precisão adequados durante a maioria das horas do dia.

Em [Akhtar e Latif 2010] os autores utilizaram a simulação discreta de eventos para modelar as funções de um *call center* e suas operações. Diferentes estratégias de roteamento de chamadas e priorização de agentes com múltiplas habilidades foram modeladas utilizando o software de simulação Witness. A verificação e validação do modelo foram realizadas durante a fase de construção do modelo. Os autores concluíram que os cenários simulados e suas suposições são úteis para os gestores tomarem decisões sobre a alocação de recursos em *call centers*.

Em [Robbins e Medeiros 2010], os autores discutem os erros associados aos resultados do modelo Erlang C quando aplicado em situações reais que ocorrem em *call centers*. Os autores buscam identificar os fatores mais influentes na produção dos erros de previsões das métricas de desempenho desse modelo. Os resultados teóricos do modelo Erlang C foram comparados com os resultados obtidos através da execução de um modelo de simulação discreta baseado em eventos de *call center*, onde várias suposições do modelo Erlang C foram relaxadas. As seguintes medidas foram comparadas entre o modelo Erlang C e o modelo de simulação: tempo de espera na fila, tempo médio para atendimento, quantidade de chamadas atendidas durante certo intervalo de tempo e taxa de abandono. Na construção do modelo de simulação foram considerados 9 fatores experimentais. Como o número de fatores era relativamente grande, foi utilizada a abordagem *Space Filling Design* baseada no *Latin Hypercube Sampling* (LHS), discutido em [Santner *et al.* 2003], com amostragem de 1.000 pontos para testar a região experimental. Para cada projeto foram executadas 500 replicações com 25 mil chamadas em cada replicação. Um período de 5.000 chamadas foi utilizado como *warm-up* para diminuir ruídos iniciais. Após todas as replicações serem executadas, os resultados foram comparados com as previsões do modelo Erlang C. O erro foi calculado como a diferença do valor teórico com o valor simulado. Os autores concluíram que o modelo

Erlang C tem maior probabilidade de ser pessimista, ou seja, o sistema real tende a ter desempenho melhor do que sua predição. No entanto, o modelo pode se tornar otimista quando a utilização do sistema é alta e as taxas de chegadas são incertas. Os erros de predição estão fortemente correlacionados com a taxa de abandono de chamadas. O modelo é mais preciso quando o número de agentes é grande e a taxa de utilização é baixa. Ficou claro que atenção especial deve ser tomada antes de usar o modelo Erlang C para fazer qualquer cálculo que necessite de precisão elevada.

Em [Chromy *et al.* 2011] os autores comparam os resultados dos cálculos de parâmetros importantes de *call centers* usando a formula Erlang C com resultados obtidos através de simulação. Eles concluíram que em situações onde tem somente um grupo de serviço e todos os agentes de atendimento possuem o mesmo desempenho, o modelo Erlang C oferece bons resultados. No entanto, apesar de ser possível utilizar o modelo Erlang C para *call centers* com vários grupos de serviços e com agentes com diferentes níveis de desempenho, os autores salientam que sua utilização torna-se mais complicada.

Em [Ibrahim *et al.* 2012], os autores comparam a precisão de modelos de predição distribucionais e de estimativas pontuais, quando aplicados em dados de *call centers*. Os autores consideraram as chamadas de 275 dias de um *call center* real, ocorridas durante o horário comercial. Essas chamadas foram divididas em períodos de 30 minutos. A transformação por meio da raiz quadrada foi utilizada para estabilizar a variância dos dados e o erro quadrático médio, objetivando comparar a precisão dos modelos. Foram geradas previsões para 85 dias com *lead time* de 1 dia e 2 semanas. Os modelos eram estimados a cada previsão diária. Os modelos distribucionais tiveram melhores resultados para prazos de previsões de 1 e 14 dias.

Em [Liao *et al.* 2012], os autores desenvolveram um modelo de *call center* de turno único com dois tipos de serviços: atendimento de chamadas telefônicas (serviço que deve ser executado o mais rápido possível) e serviços de *back-office* (o atendimento pode sofrer atraso) com foco na otimização dos recursos para diminuição de custos. Eles modelaram o processo de chegada de chamadas usando um modelo estocástico duplo não estacionário com taxas médias de chegadas aleatórias. Os autores concluíram que a combinação dos dois tipos de serviços ofereceu flexibilidade, permitindo a absorção parcial dos efeitos indesejáveis da incerteza do processo de chegadas de chamadas.

Em [Ibrahim e L'Ecuyer 2013], os autores comparam a precisão de previsões fornecidas por modelos que tentam representar o processo de chegada de chamadas em *call centers*. Os seguintes modelos foram avaliados: i) modelo de efeitos fixos (não explora

nenhuma dependência entre os dados; *ii*) modelo de efeitos mistos (leva em conta as correlações *interdays* e *intradays*); e *iii*) dois modelos de efeito misto bivariados, que exploram correlações entre diferentes tipos de chamadas. Para comparar a precisão dos modelos foram utilizados três conjuntos de dados de chamadas de um *call center* real. A transformação da raiz quadrada foi utilizada nos dados para estabilizar a variância. Foram considerados vários *lead times*, variando de horas até semanas. Os autores concluíram que os modelos que consideram diferentes estruturas de correlações nos dados obtiveram melhor precisão de predição de chegadas de chamadas.

Em [Steinmann e Freitas Filho 2013], os autores utilizam a simulação para gerar dados sintéticos para serem utilizados na avaliação de algoritmos de predição de chamadas de *call centers*. Foi construído um modelo de simulação que permite a inclusão de fatores aleatórios nos dados com diferentes tipos de impactos, como geralmente ocorre nos ambientes reais. Os autores concluíram que a geração sintética de dados por meio de simulação é uma boa alternativa para auxiliar no desenvolvimento e testes de modelos de operações de *call centers*.

Em [Sencer e Ozel 2013], os autores desenvolveram um sistema DSS (*Decision Support System*), baseado em simulação, para dar apoio à tomada de decisão sobre a força de trabalho necessária para atender demandas de chamadas telefônicas dentro dos acordos de níveis de serviços estabelecidos com os clientes. O sistema foi desenvolvido para ser utilizado em um dos maiores *call center* da Turquia. As interfaces gráficas foram criadas com o objetivo de aumentar a usabilidade e eficácia do sistema. O teste de aderência Qui-Quadrado foi utilizado para testar a aderência de distribuições teóricas de probabilidade aos dados. Na análise do processo de chegada das chamadas, por períodos de 15 minutos, verificou-se que o processo de chegada era Poisson. Como a amostra era muito grande (por volta de 30.000 chamadas) e nenhuma distribuição teórica aderiu aos dados relativos à duração das chamadas, os autores utilizaram a distribuição empírica para representar a duração da chamada no modelo de simulação. Os autores concluíram que as vantagens da simulação aliada à flexibilidade e ambiente amigável do ambiente desenvolvido fornece a possibilidade de fazer o planejamento de *workforce* mais efetivo e eficaz em *call centers*.

CAPÍTULO 3 – SISTEMA MODELADO NO ESTUDO DE CASO

3.1 – Introdução

Nesta pesquisa o sistema de telefonia da Universidade Federal de Uberlândia (UFU) foi utilizado como estudo de caso. Esse sistema permite os seguintes tipos de chamadas: *i*) ligações entre os ramais da instituição; *ii*) ligações locais; *iii*) ligações interurbanas (DDD); *iv*) ligações para celular; e *v*) ligações internacionais (DDI). Para cada terminal telefônico é associada uma categoria, a qual define o tipo de ligação que pode ser realizado através deste ramal.

A UFU possui contrato com uma operadora de telefonia fixa para completar as ligações locais, interurbanas, celulares e internacionais. Além disso, pelo fato de a UFU ser uma instituição integrante do serviço fone@RNP [RNP 2014], ela pode completar ligações interurbanas sem custo para cidades onde há cobertura do serviço.

O serviço fone@RNP é mantido pela Rede Nacional de Pesquisa (RNP). Este permite o encaminhamento de chamadas entre instituições clientes da RNP utilizando tecnologia VoIP (Voz sobre IP). Por meio desse serviço, potencialmente todos os usuários das instituições clientes que compõem essa rede conseguem se comunicar por voz (via telefone comum, telefone IP ou *softphone*) pela Internet [RNP 2014].

O fone@RNP é um serviço notoriamente conhecido pelo seu potencial de economia nas ligações interurbanas, sobretudo quando a instituição de destino completa chamadas para a rede pública de telefonia (PSTN). Porém, mais do que isso, o fone@RNP permite que o usuário final faça uso da telefonia convencional por meio de uma mobilidade própria dos serviços de Internet [RNP 2014].

A interligação do sistema de telefonia convencional da instituição com o fone@RNP geralmente é realizada utilizando um *gateway* de voz. O papel do *gateway* é converter os protocolos da telefonia convencional para os protocolos utilizados na telefonia VoIP e vice-versa.

A seguir serão apresentados os componentes do sistema de telefonia da UFU que foi investigado, bem como uma breve descrição dos seus softwares e protocolos.

3.2 – Arquitetura do Sistema de Telefonia da UFU

O sistema de telefonia utilizado nos trabalhos experimentais desta pesquisa é composto por um PABX (*Private Automatic Branch Exchange*), o qual conecta os ramais internos da instituição e faz conexão com a rede pública de telefonia (PSTN), e por um

gateway de voz que faz a mediação entre a telefonia convencional e o provedor de telefonia VoIP (fone@RNP). A conexão entre o PABX e o *gateway* de voz é realizada através de um tronco E1 com sinalização ISDN, suportando até 30 canais de voz simultâneos. Os protocolos utilizados pelo provedor VoIP são o SIP e o RTP, descritos na Seção 2.3.1 e Seção 2.3.2, respectivamente. A Figura 3.1 apresenta uma visão geral deste sistema.

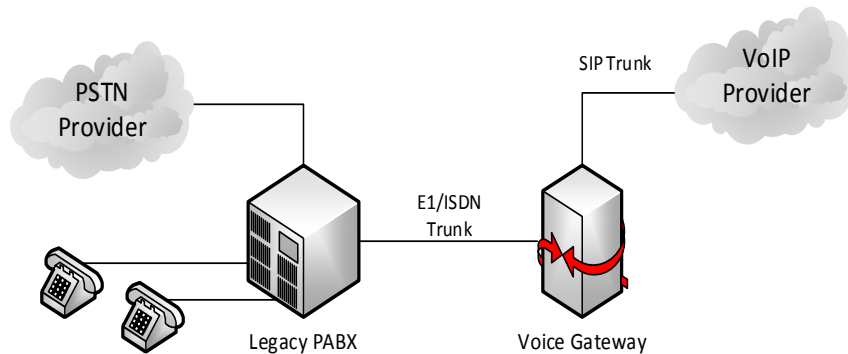


Figura 3.1 - Visão geral da atual arquitetura do sistema de telefonia da UFU.

A tabela de roteamento do *gateway* de voz contém os códigos DDD e prefixos que podem ser completados utilizando o serviço fone@RNP. Essa tabela é uma cópia local, somente de leitura, da tabela original que fica armazenada em um servidor de banco de dados SQL no provedor VoIP.

O PABX não tem capacidade de buscar rotas em banco de dados SQL. Por isso, sua lógica de roteamento é programada para encaminhar todas as chamadas DDD para o *gateway* de voz. Ao receber a chamada, o *gateway* consulta sua tabela de rotas para verificar se a chamada pode ser completada utilizando o provedor VoIP. Caso seja possível, o *gateway* encaminha a chamada para o provedor VoIP utilizando a Internet. Caso contrário, a chamada é encaminhada de volta para o PABX para ser completada utilizando o provedor PSTN. Na configuração atual, quando a chamada passa pelo *gateway* e depois retorna para o PABX, são necessários dois canais do tronco E1. Ou seja, para uma única chamada são alocados dois canais.

Além de interconectar o PABX com a telefonia VoIP, os responsáveis pelo serviço de telefonia da UFU planejam, para um futuro próximo, utilizar o *gateway* de voz para completar ligações também para celular, utilizando cartões GSM instalados em placas PCI GSM no *gateway* de voz. A Figura 3.2 ilustra o sistema de telefonia da UFU com essa funcionalidade adicionada.

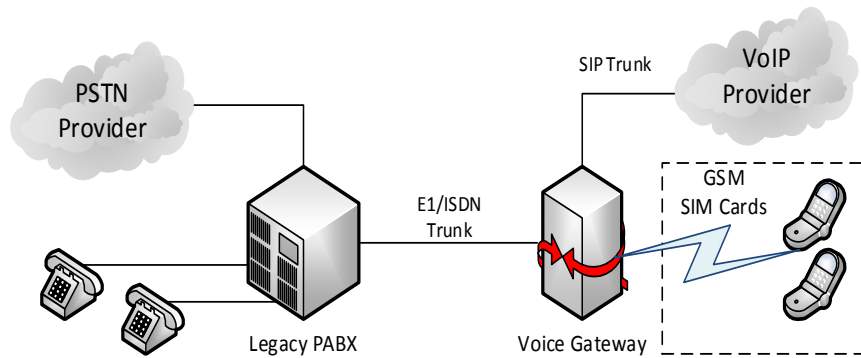


Figura 3.2 - Visão geral do sistema de telefonia da UFU com cartões GSM.

3.2.1 – PABX

PABX é um centro de distribuição telefônica que, além de permitir as ligações entre telefones internos de uma organização, se interconecta com redes de telefonia pública (PSTN) para permitir o recebimento e realização de ligações externas.

Geralmente o PABX oferece recursos adicionais, tais como:

- Transferência de chamadas;
- URA (Unidade de Resposta Audível);
- Encaminhamento de chamadas quando o ramal está ocupado;
- Puxar chamada de um grupo;
- Grupo de chamadas;
- Chamada em espera;
- Registro das chamadas realizadas.

O PABX da UFU que foi utilizado nesta pesquisa é da marca Philips, modelo SophoIS 3000 [Philips 2014]. Atualmente, este PABX suporta 1200 ramais. A Figura 3.3 mostra as placas e conexões do PABX da UFU.



Figura 3.3 - Visão frontal do PABX da UFU.

3.2.2 – Voice Gateway

O *gateway* de voz é um elemento da rede de telefonia que faz a conversão de protocolos de redes de telefonia diferentes, permitindo a interconexão entre elas. Um *gateway* de voz comum é aquele que converte os protocolos da telefonia convencional para os protocolos utilizados na telefonia VoIP e vice-versa.

A utilização do *gateway* de voz é uma solução adequada para uma organização começar a utilizar serviços VoIP, pois não exige que toda a infraestrutura da telefonia convencional seja alterada para VoIP. Ou seja, ele é adequado para fazer o processo de transição da telefonia convencional para a telefonia VoIP.

O *gateway* se conecta ao sistema de PABX legado usando troncos analógicos ou digitais. Quando são utilizados troncos analógicos, o *gateway* faz a conversão entre os protocolos VoIP e as linhas telefônicas analógicas. Já quando são utilizados os troncos digitais, a conversão é feita entre os protocolos VoIP e os protocolos utilizados pela telefonia digital (ex. E1 e T1).

Os protocolos E1 e T1 são padrões inicialmente desenvolvidos para transportar conversas de voz simultaneamente utilizando a multiplexação por divisão do tempo. O E1 é utilizado como padrão no Brasil e na Europa. Ele possui taxa de transferência de 2Mbps e pode ser dividido em 32 canais de 64kbs, com 2 canais reservados para sinalização e sincronização. Já o T1 é utilizado como padrão nos Estados Unidos. Ele possui taxa de transferência de 1544Mbps e pode ser dividido em 24 canais de 64kbs.

A capacidade de roteamento do *gateway* permite que a melhor rota para determinado destino seja escolhida para completar a chamada. Assim, caso o *gateway* esteja conectado a diferentes provedores, o administrador pode criar rotas para que a chamada seja encaminhada para o provedor que oferece menores tarifas para um dado destino.

Na UFU, a função do *gateway* de voz é realizada pelo software Asterisk [Digium 2014]. O Asterisk é um software livre, patrocinado pela empresa Digium, que pode ser utilizado para construir aplicações para comunicação. Ele é capaz de atuar como um sistema de PABX IP, *gateway* de voz, servidor de conferências e outras soluções customizadas [Digium 2014].

A conexão entre o PABX e o *gateway* de voz na UFU é realizada utilizando troncos digitais com sinalização E1/ISDN. Como o PABX e o *gateway* de voz estão instalados fisicamente em prédios distintos, uma fibra monomodo ligada em dois *modems* é utilizada para fazer a conexão física entre eles. A Figura 3.4 ilustra essa conexão.

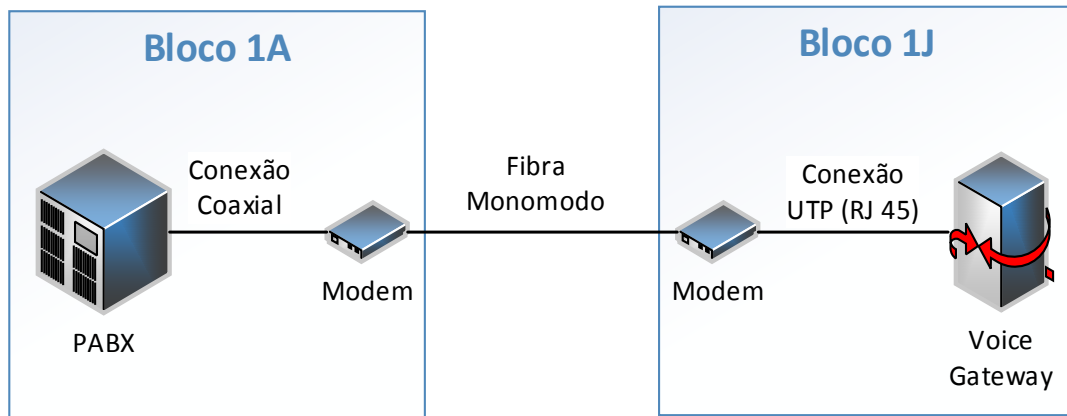


Figura 3.4 - Conexão física do PABX com o gateway de voz utilizando dois *modems*.

Os *modems* utilizados são da marca Digitel (modelo MetroFiber). Eles possuem suporte para 2 troncos E1 e possibilidade de fazer a ligação com placas E1 através de cabos UTP (com conectores RJ-45) ou cabos coaxiais (com conectores BNC). A Figura 3.5 mostra a parte frontal do *modem*, a Figura 3.6 mostra a parte traseira do *modem* que é ligado ao servidor Asterisk e a Figura 3.7 mostra a parte traseira do *modem* que é ligado ao PABX.



Figura 3.5 - Visão frontal do *modem*.



Figura 3.6 - Visão da parte traseira do *modem* que é ligado ao servidor Asterisk (cabo UTP com conector RJ-45).



Figura 3.7 - Visão da parte traseira do *modem* que é ligado ao PABX (cabo coaxial com conector BNC).

A placa de telefonia E1 utilizada no servidor Asterisk é da marca Digium (modelo TE121P), com barramento PCIe e saída RJ-45. Assim, como os *modems* Digitel aceitam conexões através de cabos UTP e conectores RJ-45, não foi necessário utilizar *baloons* para fazer a conexão entre o *modem* e o servidor.

A Figura 3.8 mostra a parte frontal do servidor físico onde está instalada a máquina virtual que executa o software Asterisk. Já a Figura 3.9 mostra a parte traseira do servidor. O cabo UTP de cor amarela, que aparece na Figura 3.9, faz a ligação do *modem* com a placa PCI-Express do servidor. Os conectores do cabo UTP são RJ-45.



Figura 3.8 - Visão frontal do Servidor XVM-VOIP.



Figura 3.9 - Visão traseira do servidor XVM-VOIP, mostrando a ligação do cabo com conector RJ45 vindo do *modem* para a placa PCI do servidor.

Para configurar a placa Digium E1 no software Asterisk, utilizou-se os parâmetros apresentados pela Figura 3.10, os quais estão armazenados no arquivo de configuração *chan_dahdi.conf* do Asterisk.

```
[channels]
language=en
context=from-pabx
switchtype=euroisdn
signaling=pri_net
overlapdial=yes

context=from-pabx
group=1
channel=>1-15,17-31
```

Figura 3.10 - Configuração do canal DAHDI no Asterisk.

Como o software Asterisk foi instalado em uma máquina virtual no XEN, foi necessário realizar alguns procedimentos para que a placa Digium E1 PCIe pudesse ser utilizada pela máquina virtual. Os passos listados na Figura 3.11 foram executados no servidor XVM-VOIP, que possui a versão 6.2 da plataforma XenServer.

1 - Descobrir o ID da placa Digium no servidor XEN Server:

```
# lspsci
```

2 - Em /boot/extlinux.conf, adicionar a opção pciback.hide (id_da placa) no parâmetro "append"

```
append /boot/xen.gz mem=1024G .... pciback.hide=(id_da_placa)
```

3 - Com a máquina virtual desligada, configurar o parâmetro other-config da máquina virtual Asterisk

```
xe vm-param-set other-config:pci=0/0000:id_da_placa uuid=UID_DA_VM
```

Figura 3.11 - Configuração da placa E1 PCIe no servidor XVM-VOIP.

3.2.3 – CDR

CDR (*Call Detail Record*) é o termo padrão para o registro de uma conexão telefônica, o qual contém os detalhes da chamada que utilizou esta conexão. Os equipamentos de telefonia (PABX e *gateways* de voz) geram e armazenam os CDRs para que esses dados sejam posteriormente utilizados por ferramentas especializadas na construção de relatórios e também na realização do processo de bilhetagem e de faturamento. O CDR também pode ser utilizado pelos administradores do sistema como ferramenta de identificação de problemas.

Os CDRs utilizados nessa pesquisa foram coletados no PABX e no *gateway* de voz.

3.2.3.1 – CDR do Gateway de voz

O software Asterisk permite a escolha do serviço/tipo de armazenamento para os registros das chamadas. A lista abaixo apresenta alguns dos serviços/tipos suportados.

- ODBC
- CSV
- Mysql
- Postgres
- Sqlite
- Microsoft SQL Server

Ao decidir pela utilização de um destes serviços, deve ser realizada a devida configuração do módulo correspondente.

No *gateway* de voz da UFU, utiliza-se o banco de dados Postgres para armazenar os registros das chamadas. A Tabela 3.1 descreve o significado dos campos do CDR.

Tabela 3.1 - Campos de um CDR no *gateway* de voz.

Nome	Descrição
Calldate	Data e hora do início da conexão
Clid	<i>Caller ID</i> completo do chamador
Src	<i>Caller ID Number</i> do chamador
Dst	Extensão destino da chamada
Dcontext	Contexto de destino da chamada
Channel	Canal do chamador
Dstchannel	Canal chamado
Lastapp	Última aplicação do plano de discagem que foi executada
Lastdata	Argumentos passados para a última aplicação
Duration	Número de segundos entre o início e o fim da chamada
Billsec	Número de segundos entre a resposta da chamada e o fim da chamada
Disposition	Indicador do que aconteceu com a chamada. Pode ter os valores: NO ANSWER, FAILED, BUSY, ANSWERED ou UNKNOWN.
Amaflags	<i>Flag Automatic Message Accounting</i> (AMA) associada com esta chamada.
Accountcode	Código associado com a chamada (Pode ser utilizado para fins de tarifação)
Uniqueid	Identificador único para o canal de origem
Userfield	Campo de propósito geral

A Tabela 3.2 apresenta um registro detalhado de uma chamada completada (*Disposition=ANSWERED*) utilizando o provedor PSTN. A chamada foi realizada a partir do ramal "4297" do PABX e o número de destino é de Brasília (012-61-33638636). O valor 099 do campo *Accountcode* indica que a chamada foi realizada utilizando o provedor PSTN. Note

que a duração da chamada foi de 38 segundos. No entanto, para a bilhetagem foi considerado somente o tempo de fala (22 segundos).

Tabela 3.2 - CDR de uma chamada atendida completada via provedor PSTN.

Nome	Descrição
Calldate	2012-09-25 14:17:20
Clid	4297
Src	4297
Dst	0126133638636
Dcontext	from-pabx
Channel	DAHDI/i1/4297-2826
Dstchannel	DAHDI/i1/*80126133638636-2828
Lastapp	Dial
Lastdata	DAHDI/g1/*80126133638636,60,rtT
Duration	38
Billsec	22
Disposition	ANSWERED
Amaflags	3
Accountcode	099
Uniqueid	1348593440.25278
Userfield	-

A Tabela 3.3 apresenta um registro detalhado de uma chamada completada utilizando o provedor VoIP. A chamada foi realizada a partir do ramal "4297" do PABX e o número de destino é de Belo Horizonte (012-31-34095829). O valor 001 para o *Accountcode* indica que a chamada foi realizada utilizando o provedor VoIP. A duração da chamada foi de 132 segundos e o tempo de fala 121 segundos.

Tabela 3.3 - CDR de uma chamada atendida encaminhada para o provedor VoIP.

Nome	Descrição
Calldate	2012-10-05 09:05:35
Clid	4297
Src	4297
Dst	0123134095829
Dcontext	from-pabx
Channel	DAHDI/i1/4297-66
Dstchannel	SIP/voip.ufu.br-0000003d
Lastapp	Dial
Lastdata	SIP/553134095829@voip.ufu.br,60,rtT
Duration	132
Billsec	121
Disposition	ANSWERED
Amaflags	3
Accountcode	001
Uniqueid	1349438735.162
Userfield	-

Além das chamadas atendidas (*Disposition=Answered*), também foram consideradas as chamadas DDD que não foram atendidas (*Disposition=No Answer*). Estas chamadas foram consideradas porque, mesmo não atendidas, há utilização de recursos quando a chamada está em estado *ringing* (estado em que o telefone chamado está tocando e ainda não foi atendido).

A Tabela 3.4 apresenta um registro detalhado de uma chamada que foi encaminhada para o provedor PSTN (*Accountcode=099*) e não foi atendida. Note que a duração da chamada foi de 51 segundos. Durante esse tempo, apesar de a ligação não ter sido atendida, houve a utilização de 2 canais do tronco E1.

Tabela 3.4 - CDR de uma chamada não atendida encaminhada para o provedor PSTN.

Nome	Descrição
Calldate	2012-01-24 16:20:02
Clid	4978
Src	4978
Dst	0126434425777
Dcontext	from-pabx
Channel	DAHDI/i1/4978-3525
Dstchannel	DAHDI/i1/*80126434425777-3526
Lastapp	Dial
Lastdata	DAHDI/g1/*80126434425777,60,rtT
Duration	51
Billsec	0
Disposition	NO ANSWER
Amaflags	3
Accountcode	099
Uniqueid	1327429202.32671
Userfield	-

3.2.3.2 – CDR do PABX

Os registros detalhados das chamadas para celular foram obtidos com os responsáveis pelo sistema PABX. Como as chamadas para celular ainda não passam pelo *gateway* de voz, as informações sobre essas chamadas foram limitadas a informações que foram fornecidas pelos responsáveis pelo PABX.

A Tabela 3.5 apresenta os campos do CDR das chamadas para celular que foram disponibilizados para esta pesquisa. Neste caso, foram cedidos somente os registros das chamadas que foram atendidas. Além disso, a duração total da chamada não foi disponibilizada. O tempo disponibilizado corresponde somente ao tempo de conversa durante a chamada. Os dados foram disponibilizados em um arquivo de formato texto.

Tabela 3.5 - Campos de um CDR de chamadas para celular coletado no PABX.

Nome do Campo	Função
Date	Data do início da ligação
Time	Horário do início da ligação
Billsec	Duração da conversa

Neste estudo, os dados disponibilizados foram carregados para uma tabela em um banco de dados Postgres para facilitar as consultas realizadas com o objetivo de analisar estatisticamente os dados.

3.3 – Considerações Finais

Neste capítulo foram apresentados os detalhes do ambiente que foi investigado durante a pesquisa.

Os dados utilizados nesta pesquisa foram obtidos dos CDRs dos sistemas *gateway* de voz e PABX da UFU. O entendimento dos campos dos CDRs foi importante para fazer a modelagem precisa das chamadas. Na fase de coleta de dados dos CDRs do *gateway* de voz, o campo *Accountcode* foi utilizado para diferenciar as chamadas do tipo I (VoIP) e tipo II (PSTN). Além disso, o campo *Disposition* foi utilizado para filtrar as chamadas atendidas e não atendidas (*Disposition=ANSWERED* e *Disposition=NO ANSWER*).

O padrão utilizado para fazer a conexão entre o *gateway* de voz e o PABX é o E1. Assim, o tronco E1 entre o PABX e o *gateway* permite até 30 chamadas simultâneas.

No próximo capítulo serão apresentados os métodos e materiais utilizados para modelar o sistema descrito neste capítulo.

CAPÍTULO 4 – MODELAGEM E SIMULAÇÃO DE UM GATEWAY DE VOZ

4.1 – Introdução

Este capítulo apresenta os métodos e materiais utilizados para modelar e simular o sistema de telefonia detalhado no Capítulo 3, com o objetivo de planejar a capacidade dos recursos de um *gateway* de voz.

4.2 – Modelo de Simulação

A Figura 4.1 apresenta uma visão geral do modelo de simulação criado para representar o sistema investigado apresentado no Capítulo 3.

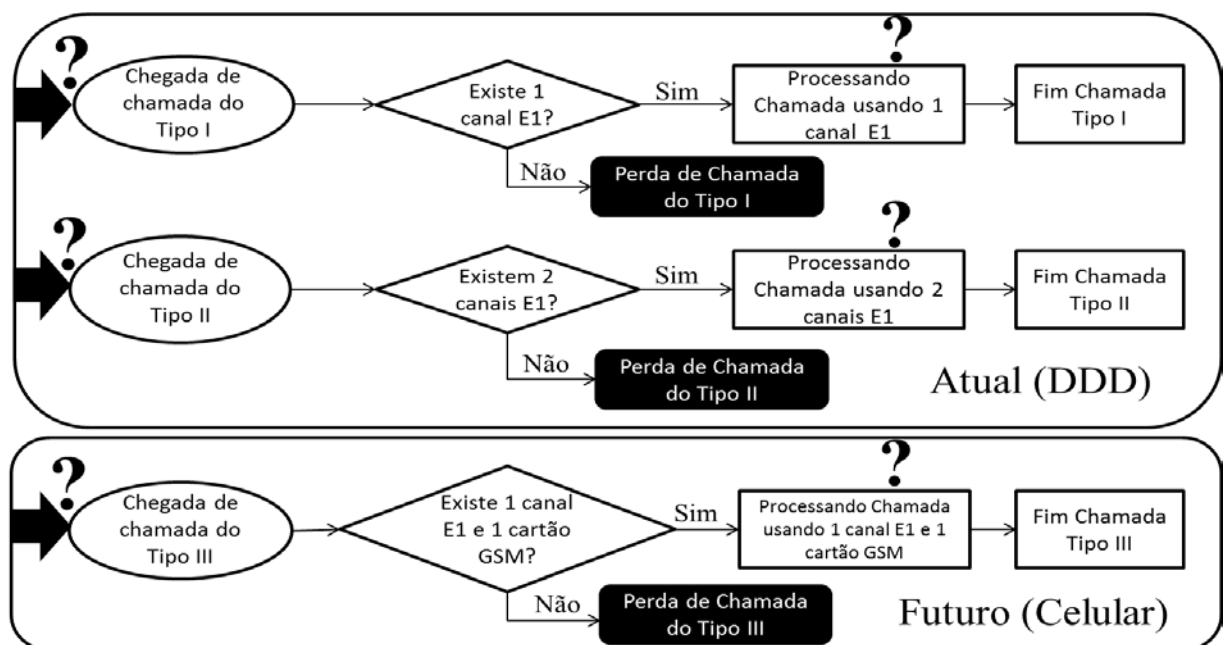


Figura 4.1 - Visão geral do modelo de simulação.

Como o objetivo é planejar a capacidade dos recursos utilizados pelo *gateway* de voz, foram consideradas as chamadas DDD (que no sistema atual já usam os recursos do *gateway*) e as chamadas para celular (que utilizarão os recursos do *gateway* no futuro). Essas chamadas foram categorizadas em três tipos:

- (I) - Chamadas DDD completadas utilizando um provedor VoIP;
- (II) - Chamadas DDD completadas utilizando um provedor PSTN;
- (III) - Chamadas para celular completadas utilizando um provedor GSM.

Essa categorização em três tipos foi necessária porque cada um dos tipos de chamadas necessita de recursos diferentes para ser completada. Portanto, cada um dos tipos de chamadas foi modelado separadamente.

Os instantes das chegadas de chamadas no *gateway* são determinados por tempos aleatórios que são gerados com base em distribuições de probabilidade que representam a população (tempo entre chegadas) de cada um dos três tipos de chamadas considerados.

Assim que uma chamada chega ao *gateway* de voz, é verificado se existem recursos suficientes para completá-la. Em caso positivo, os recursos (canais E1 e/ou cartões GSM) são alocados durante a duração da chamada. Os recursos utilizados durante a ligação são liberados ao final da chamada. Caso não existam recursos suficientes para completá-la, ela é perdida.

Os recursos que foram analisados no *gateway* são os canais E1 e os cartões GSM. A Tabela 4.1 apresenta a quantidade de recursos necessários para processar cada tipo de chamada.

Tabela 4.1 - Recursos do *gateway* utilizados por cada tipo de chamada.

Tipo	Recursos
Tipo I (VoIP)	1 canal (E1)
Tipo II (PSTN)	2 canais (E1)
Tipo III (GSM)	1 canal (E1) e 1 cartão GSM

Os tempos em que os recursos são utilizados durante a chamada são determinados por tempos aleatórios que são gerados com base em distribuições de probabilidade que representam a população (duração da chamada) de cada um dos três tipos de chamadas considerados.

Os pontos de interrogação na Figura 4.1 representam as distribuições de probabilidade que serão utilizadas para representar o tempo entre chegada e as durações dos três tipos de chamadas consideradas. O processo de identificação destas distribuições será explicado nas seções seguintes.

4.3 – Coleta de Dados

Duas variáveis aleatórias foram modeladas para servir de entrada para o modelo: "Tempo entre Chegada de Chamadas" e "Duração da Chamada". A variável "Tempo entre Chegada de Chamadas" foi utilizada para modelar o comportamento das chegadas de chamadas no *gateway*. Já a variável "Duração da Chamada" foi utilizada para modelar a quantidade de tempo que os recursos são utilizados durante o processamento da chamada. Ambas as variáveis foram modeladas separadamente para os três tipos de chamadas.

Os dados para modelar essas duas variáveis foram obtidos a partir dos CDR coletados no PABX e no *gateway* de voz do sistema investigado. Como explicado no Capítulo 3, os dados das chamadas do tipo I e II foram coletados no *gateway* de voz e os dados das chamadas do tipo III foram coletados no PABX.

No processo de amostragem, foram considerados os dados de chamadas realizadas no período comercial (08:00 às 18:00) de Jan/2012 a Dez/2012. Como o gargalo do sistema ocorre quando a taxa de utilização dos recursos é alta, a partir da amostra anual foram selecionados os 30 dias com o maior volume de ligações. Não faz sentido testar o modelo quando o volume de ligações é muito baixo, pois neste caso a taxa de utilização dos recursos não representa problemas de desempenho. Além disso, os períodos da manhã (08:00 às 13:00) e tarde (13:00 às 18:00) foram simulados separadamente para fornecer informações adicionais aos analistas sobre a diferença de utilização dos recursos nestes dois períodos. As estatísticas descritivas do número de chamadas e da duração das chamadas dos 30 dias considerados são apresentadas na Tabela 4.2 e na Tabela 4.3, respectivamente.

Tabela 4.2 - Estatísticas descritivas para o número de chamadas.

	Tipo I (VoIP)	Tipo II (PSTN)	Tipo III (GSM)
Total	11921	19112	46557
Média	397,4	637,1	1551,9
Mediana	378	627	1552
Mínimo	232	539	1418
Máximo	540	780	1748
Desvio Padrão	74,8	56,5	76,98

Tabela 4.3 - Estatísticas descritivas para a duração das chamadas (em segundos).

	Tipo I (VoIP)	Tipo II (PSTN)	Tipo III (GSM)
Média	118,9	130	126,4
Mediana	52	69	78
Mínimo	4	4	6
Máximo	5162	3843	4461
Desvio Padrão	200,7	195,2	164,8

Para as ligações dos tipos I e II, foram consideradas as chamadas completadas e não completadas. Também, para esses dois tipos de chamadas, foram considerados os tempos em que as chamadas estavam no estado *ringing* (estado em que o telefone chamado está tocando e a chamada ainda não foi atendida), pois durante este tempo ocorre utilização de 1 canal (E1) para as ligações do tipo I e 2 canais (E1) para as chamadas do tipo II. Devido à forma como foi configurado o sistema responsável por gerar os CDRs das ligações no sistema PABX, só foi possível utilizar os dados das chamadas do tipo III que foram completadas. Foi recomendado aos analistas responsáveis pelo sistema PABX habilitar a opção para registrar também as informações das chamadas que não são atendidas para que elas possam ser consideradas nas pesquisas futuras.

Nas seções 4.4 e 4.5 serão apresentados os detalhes da modelagem das variáveis "Tempo entre Chegada de Chamadas" e "Duração da Chamada".

4.4 – Modelagem da Variável "Tempo entre Chegada de Chamadas"

Considerando as amostras completas, os três testes de aderência que foram utilizados neste trabalho (A-D, K-S e χ^2) rejeitaram a hipótese de que os dados das variáveis aleatórias que representam o tempo entre chegadas, dos três tipos de chamadas, seguissem alguma distribuição de probabilidade teórica conhecida.

Devido à alta variabilidade da taxa média de chegada de chamadas ao longo do dia, foi utilizada a proposta descrita em [Law e Kelton 1991], a qual recomenda para estes casos agrupar os dados em períodos menores para depois realizar os testes de aderência. Desse modo, foi utilizado o agrupamento de chamadas em períodos de 15 minutos. Caso os testes de aderência utilizando os dados agrupados por períodos de 15 minutos rejeitassem a hipótese de que os dados seguissem alguma distribuição teórica de probabilidade, o agrupamento deveria ser feito utilizando períodos menores.

A Tabela 4.4 apresenta os valores das taxas médias de chegadas de chamadas para cada período de 15 minutos. A Figura 4.2 apresenta os dados de forma gráfica.

Tabela 4.4 - Taxa média de chegada de chamadas dos 30 dias com maior volume de chamadas agrupadas em períodos de 15 minutos.

Período	Valor			Período	Valor		
	I	II	III		I	II	III
08:00-08:15	3,9	6,03	18,53	13:00-13:15	2,9	7,63	24,76
08:15-08:30	6,33	10,06	26,16	13:15-13:30	3,5	10,43	32,26
08:30-08:45	8,13	13,2	31,96	13:30-13:45	6,73	11,73	36,73
08:45-09:00	10,3	15,13	37,26	13:45-14:00	7,56	16,3	40,83
09:00-09:15	10,46	18,5	40,66	14:00-14:15	12,33	19,3	44,93
09:15-09:30	12,36	19,73	42,7	14:15-14:30	13,96	19,1	43,9
09:30-09:45	14,86	20,46	43,86	14:30-14:45	15,43	21,93	44,7
09:45-10:00	16,23	20,7	48	14:45-15:00	13,6	20,73	42,26
10:00-10:15	14,23	22,9	45,5	15:00-15:15	15,66	18,96	45,1
10:15-10:30	15,76	22,33	47,16	15:15-15:30	13,1	18	43,53
10:30-10:45	16,33	23,93	46,43	15:30-15:45	12,4	19	44,53
10:45-11:00	13,8	22,36	51,46	15:45-16:00	14,23	17,93	43,6
11:00-11:15	12,9	20,16	51,53	16:00-16:15	12,83	18	44,26
11:15-11:30	11,53	16,76	49,33	16:15-16:30	12,3	16,2	43,9
11:30-11:45	7,13	13,93	40,26	16:30-16:45	10,36	17,2	45,5
11:45-12:00	5,56	11,7	33,93	16:45-17:00	10,43	16,33	45,6
12:00-12:15	3,86	7,63	26,9	17:00-17:15	8,06	13,16	39,03
12:15-12:30	2,13	5,13	21,63	17:15-17:30	6,86	10,93	37,56
12:30-12:45	1,43	4,76	21,5	17:30-17:45	3,3	7,73	27,93
12:45-13:00	1,83	6,5	22	17:45-18:00	2,5	5,4	20,1

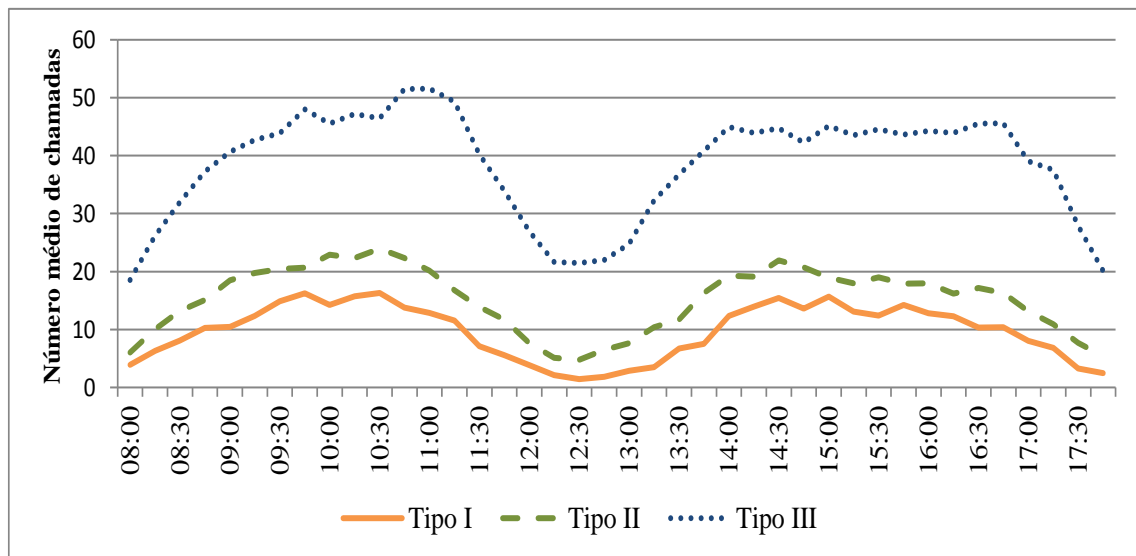


Figura 4.2 - Taxa média de chegada de chamadas por períodos de 15 minutos.

No eixo x tem-se os horários das chamadas que foram consideradas (08:00 às 18:00). No eixo y tem-se a média do número de chamadas. Cada ponto na curva representa a taxa média do número de chamadas durante o período de 15 minutos.

Nota-se que, considerando o agrupamento por períodos de 15 minutos, a taxa média de chegadas das chamadas do tipo III (GSM) é superior às taxas dos demais tipos de chamadas em todos os períodos do dia. Enquanto que as ligações do tipo I (VoIP) possuem a menor taxa de média de chamadas durante todos os períodos.

Baseado no conjunto de dados agrupados, os testes de aderência foram realizados novamente e desta vez indicaram que o novo conjunto de dados segue uma distribuição exponencial para todos os períodos do dia. Como era esperado, o valor médio varia de período para período, isto é, o processo de chegada é Poisson, mas a taxa de chegada varia ao longo do tempo. Este padrão sugere que o processo de chegada seja Poisson não estacionário, $\{N(t), t \geq 0\}$, cujas características são [Law e Kelton 1991]:

- 1 - As chegadas ocorrem uma por vez;
- 2 - $N(t + s) - N(t)$ é independente de $\{N(u), 0 \leq u \leq t\}$, ou seja, o número de chegadas em intervalos diferentes são independentes.

Assim, as variáveis aleatórias "Tempo entre Chegada de Chamadas" dos três tipos de chamadas foram modeladas usando distribuição exponencial não estacionária [Law e Kelton 1991].

4.5 – Modelagem da Variável "Duração da Chamada"

Para as chamadas do tipo I, o teste de aderência χ^2 para a variável aleatória "Duração da Chamada" indicou uma aderência aceitável para a distribuição LogNormal ($p\text{-value} = 0,0589$). Os valores estimados dos parâmetros da distribuição, *LogMean* e *LogStd*, foram 119 e 218, respectivamente.

Para os tipos de chamadas II e III, os testes de aderência χ^2 e A-D para a variável aleatória "Duração da Chamada" apresentaram $p\text{-values}$ menores que 0,05, rejeitando a hipótese de que as amostras seguem alguma das distribuições teóricas testadas. O teste K-S não foi utilizado para esses dois tipos de chamadas porque ele suporta um conjunto de dados de no máximo 8192 elementos [Arena 2014]. As Tabela 4.5 e Tabela 4.6 apresentam, respectivamente, os valores das estatísticas dos testes A-D e χ^2 para a "Duração da Chamada" dos tipos II e III. Os $p\text{-values}$ também são apresentados nas tabelas.

Tabela 4.5 - Resultados do teste A-D para a "Duração da Chamada".

Distribuição	Tipo II		Tipo III	
	AD	$p\text{-value}$	AD	$p\text{-value}$
Normal	2439,24	< 0,005	4357,29	< 0,005
Lognormal	135,66	< 0,005	23,48	< 0,005
Exponencial	582,42	< 0,003	822,43	< 0,003
Weibull	572,50	< 0,010	753,74	< 0,010
Gamma	598,47	< 0,005	667,00	< 0,005

Tabela 4.6 - Resultados do teste χ^2 para a "Duração da Chamada".

Distribuição	Tipo II		Tipo III	
	Q^2	$p\text{-value}$	Q^2	$p\text{-value}$
Normal	2,99e+004	< 0,005	4,99e+004	< 0,005
Lognormal	691	< 0,005	221	< 0,005
Exponencial	4,61e+003	< 0,005	5,23e+003	< 0,005
Weibull	6,32e+003	< 0,005	6,81e+003	< 0,005
Gamma	4,11e+003	< 0,005	7,17e+003	< 0,005

Devido ao fato de as amostras serem de tamanho muito grande (>11.000), não é incomum que os testes de aderência rejeitem a hipótese nula [Law e Kelton 1991]. Portanto, também foi utilizada a análise gráfica para verificar a aderência das amostras das chamadas dos tipos II e III a alguma distribuição teórica de probabilidade.

A Figura 4.3 mostra o gráfico de probabilidade ($p\text{-p plot}$), considerando a distribuição LogNormal, para a variável aleatória "Duração da Chamada" do tipo de chamada II (PSTN). A Figura 4.4 apresenta o mesmo gráfico para o tipo de chamada III (GSM).

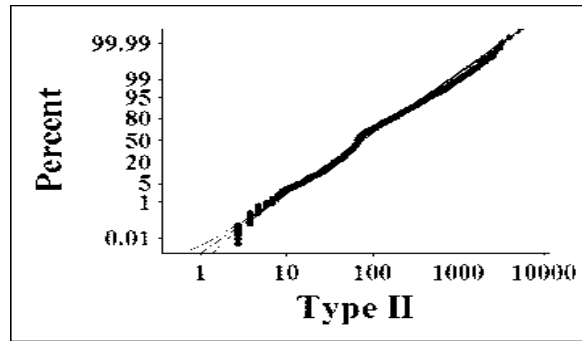


Figura 4.3 - Gráfico de probabilidade (*p-p plot*) da variável aleatória "Duração da Chamada" para o tipo de chamada II (PSTN).

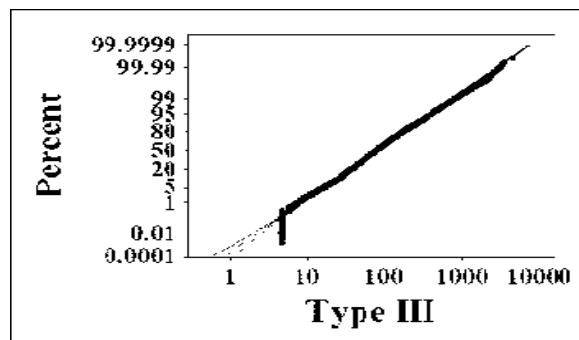


Figura 4.4 - Gráfico de probabilidade (*p-p plot*) da variável aleatória "Duração da Chamada" para o tipo de chamada III (GSM).

A análise gráfica mostra a boa aderência à distribuição LogNormal para as duas amostras, o que motivou o seu uso neste estudo. A Tabela 4.7 mostra os valores dos parâmetros *LogMean* e *LogStd* estimados para os três tipos de chamadas.

Tabela 4.7 - Parâmetros da distribuição LogNormal para a variável "Duração da chamada".

Tipo de Chamada	LogMean	LogStd
I (VoIP)	119	218
II (PSTN)	123	155
III (GSM)	123	144

4.6 – Variáveis de Resposta da Simulação

Como o objetivo da simulação é fazer o planejamento de capacidade dos canais E1 e cartões GSM instalados no *gateway* de voz, em cada simulação foram calculados os valores máximo e médio e a taxa média de utilização desses recursos, representados pelas variáveis de respostas *E1_max*, *E1_avg*, *E1_rate*, *GSM_max*, *GSM_avg* e *GSM_rate*. As variáveis *Mobile_lost* e *LDC_lost* armazenam as informações de chamadas perdidas para celular e DDD, respectivamente. Além disso, para cada tipo de chamada foi computada a quantidade de chamadas geradas (*VoIP_total*, *PSTN_total* e *Mobile_total*) para auxiliar no processo de validação do modelo e para fazer o cálculo da taxa de perda de chamadas.

A Tabela 4.8 apresenta uma breve descrição dessas variáveis de resposta que foram estimadas em cenário simulado. Além dos valores estimados, também foi computado o valor referente ao semi-intervalo de confiança de 95% ($\pm HW_{CI}$) para cada uma das variáveis.

Tabela 4.8 - Variáveis de resposta obtidas por simulação.

Variáveis de Resposta	Descrição
<i>E1_max</i>	Quantidade máxima de canais E1 utilizados simultaneamente
<i>E1_avg</i>	Quantidade média de canais E1 utilizados simultaneamente
<i>E1_rate</i>	Taxa média de utilização de canais E1
<i>GSM_max</i>	Quantidade máxima de cartões GSM utilizados simultaneamente
<i>GSM_avg</i>	Quantidade média de cartões GSM utilizados simultaneamente
<i>GSM_rate</i>	Taxa média de utilização de cartões GSM
<i>VoIP_total</i>	Número de ligações DDD utilizando o provedor VoIP (tipo I)
<i>PSTN_total</i>	Número de ligações DDD utilizando o provedor PSTN (tipo II)
<i>Mobile_total</i>	Número de ligações para celular (tipo III)
<i>LDC_lost</i>	Número de chamadas DDD perdidas (tipos II e III)
<i>Mobile_lost</i>	Número de chamadas para celular perdidas (tipo III)

4.7 – Cenários Simulados

Neste estudo, seis cenários diferentes foram simulados. O objetivo do cenário #1 é avaliar a capacidade do sistema *gateway* de voz frente à demanda atual de chamadas DDD (chamadas dos tipos I e II). Dado que no sistema real o *gateway* de voz e o sistema PABX são interligados usando um tronco E1 (suportando até 30 chamadas de voz simultaneamente), neste cenário foi considerada exatamente esta configuração atual de canais E1 do sistema real.

O objetivo do cenário #2 é fazer a avaliação da capacidade dos recursos do *gateway* de voz caso as chamadas para celular comecem a utilizá-lo para serem completadas. Portanto, além das chamadas DDD (tipos I e II), nesse cenário as chamadas para celular (tipo III) também utilizam os recursos do *gateway*. Tal cenário visa descobrir a quantidade de cartões GSM que são necessários para atender essa nova demanda de chamada para celulares. Portanto, para não haver perda de chamadas, foram considerados recursos ilimitados de cartões GSM e recursos ilimitados de canais E1.

O objetivo do cenário #3 é avaliar a capacidade dos recursos do *gateway* de voz considerando a configuração atual de 30 canais E1 e a instalação de 2 placas PCI no *gateway* de voz, com cada placa suportando 5 cartões GSM. Como o hardware atual em que o *gateway* de voz é executado suporta apenas mais 2 placas PCI, a simulação desse cenário visa descobrir se com apenas 8 cartões GSM o número de chamadas que são perdidas é aceitável. Caso o número de chamadas perdidas seja aceitável, não seria necessário investimento em um novo hardware devido à limitação de *slots* PCI no atual servidor.

O objetivo do cenário #4 é avaliar a utilização de recursos do *gateway* caso haja um incremento de demanda de chamadas no sistema. O aumento de demanda considerado das chamadas dos tipos I, II e III foi respectivamente de 3%, 5% e 10%.

No cenário #5, o modelo de simulação foi modificado para modelar uma proposta de otimização, descrita no próximo capítulo, que diminui o número de canais E1 necessários para completar as chamadas do tipo II (PSTN) de 2 canais para 1 canal. O objetivo da simulação é avaliar o valor da economia de recursos de canais E1 que a proposta de otimização poderá trazer caso seja implementada no sistema real. O valor percentual da economia será calculado baseado na comparação dos resultados deste cenário com os resultados do cenário #4, pois esses dois cenários consideram os mesmos volumes de chamadas e as mesmas configurações de recursos.

O objetivo do cenário #6 é quantificar o que o valor percentual a mais de disponibilidade dos canais E1, oferecido pela otimização apresentada no cenário #5, representa em número de chamadas a mais que podem ser completadas utilizando o *gateway* de voz.

A Tabela 4.9 resume os cenários que foram simulados neste estudo. Para cada cenário são apresentados os tipos de chamadas considerados e os recursos do *gateway* para atender a demanda considerada.

Tabela 4.9 - Tipos de chamadas simuladas e recursos utilizados.

Cenário	Tipo de chamadas	Recursos
#1	I e II.	30 canais (E1)
#2	I, II e III.	Canais ilimitados (E1) + cartões GSM ilimitados
#3	I, II e III.	30 canais (E1) + 8 cartões GSM
#4	I, II e III com incremento de demanda de 3%, 5% e 10% dos tipos I, II e III, respectivamente.	30 canais (E1) + 16 cartões GSM
#5	I, II e III com incremento de demanda de 3%, 5% e 10%, respectivamente.	30 canais (E1) + 16 cartões GSM + configuração otimizada.
#6	I, II e III com incremento necessário para manter a taxa de utilização do tronco E1 (30 canais) em 100% durante todo o tempo de simulação.	30 canais (E1) + cartões GSM ilimitados + configuração atual e configuração otimizada.

4.8 – Ferramentas Utilizadas na Modelagem e Simulação

4.8.1 – Software Arena

Para criar o modelo e simular os cenários descritos na seção anterior, utilizou-se o software Arena [Arena 2014] versão 12.0, com o tipo de licença *Training/Evaluation* (modo

student). Essa licença permite a utilização de todas as funcionalidades do software. No entanto, o tamanho do modelo que pode ser construído é restrito. Tal limitação não foi um problema para este trabalho.

O software Arena ofereceu as seguintes capacidades para o desenvolvimento e simulação do modelo:

- Construção do modelo utilizando flowcharts;
- Inclusão de diferentes distribuições de probabilidade para as variáveis aleatórias;
- Recursos de animação do modelo;
- Escolha do número de replicações e período de *warm-up*;
- Cálculo dos intervalos de confiança.

A Figura 4.5 apresenta o modelo criado utilizando o Arena.

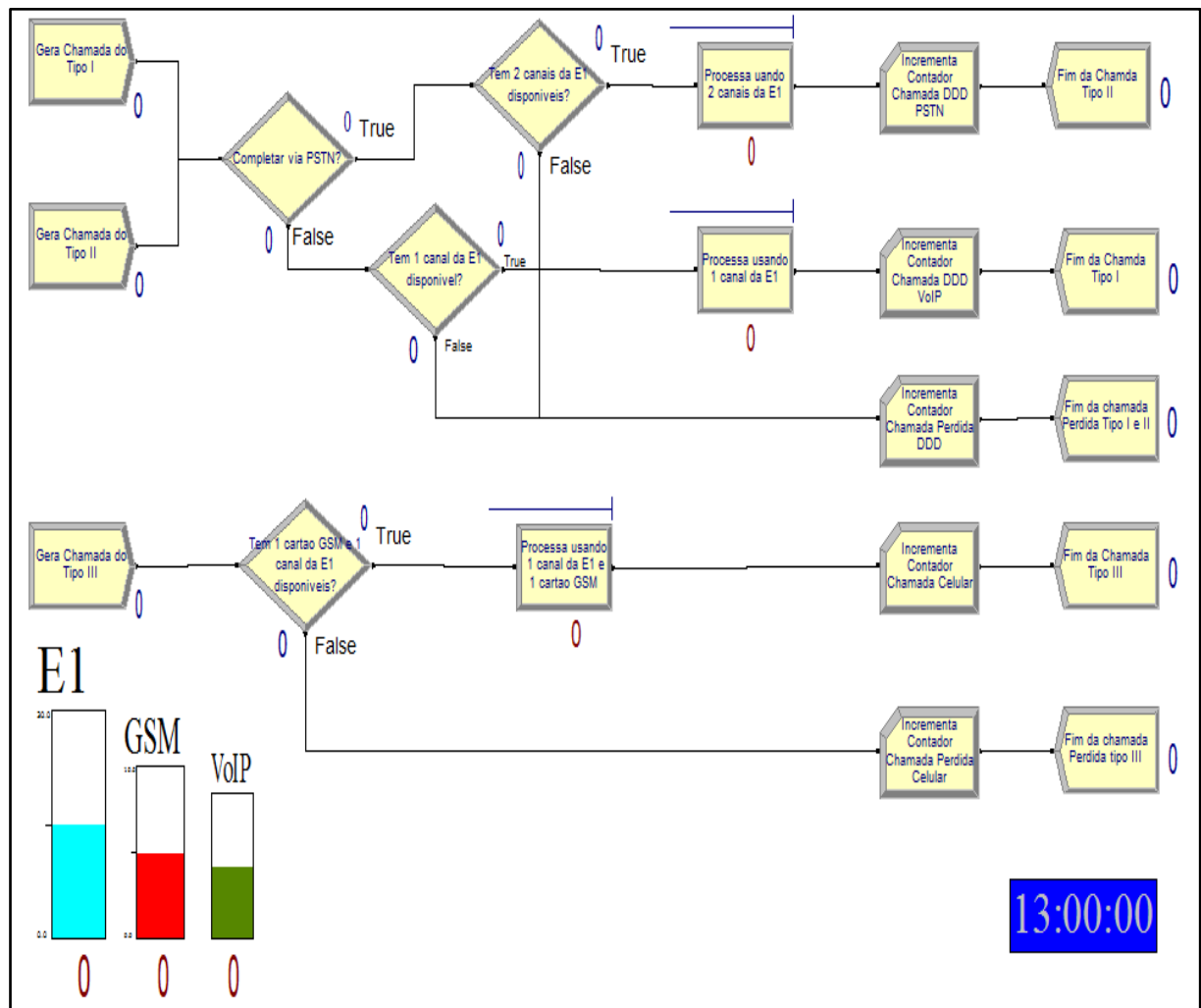


Figura 4.5 - Modelo construído no Arena.

As ligações dos tipos I, II e III foram geradas utilizando módulo *create* do ARENA. Na Figura 4.5, os módulos *create* possuem os *labels* "Gera Chamada do Tipo I", "Gera

Chamada do Tipo II" e "Gera Chamada do Tipo III". Os números "0", que aparecem ao lado desses módulos, representam as quantidades de ligações que foram geradas por eles durante a replicação corrente.

Na Figura 4.5, os losangos são os módulos de decisão (*decide*) do Arena. Estes módulos foram utilizados para direcionar a chamada com base em determinada condição (por exemplo, se existem recursos suficientes para completar a chamada).

O tempo de processamento da chamada e a quantidade de recursos utilizados pela chamada foram configurados utilizando o módulo *process* do Arena, simbolizado pelo retângulo.

O módulo *record* é utilizado para armazenar o valor de um indicador de desempenho. Os cinco módulos *record* utilizados no modelo possuem os *labels* "Incrementa Contador Chamada DDD PSTN", "Incrementa Contador Chamada DDD VoIP", "Incrementa Contador Chamada Celular", "Incrementa Contador Chamada Perdida DDD" e "Incrementa Contador Chamada Perdida Celular". Eles foram utilizados para armazenar a quantidade de chamadas completadas e perdidas.

Os gráficos de barra do Arena podem ser utilizados para permitir a visualização em tempo real de algumas informações quantitativas durante a replicação do modelo. No modelo apresentado pela Figura 4.5, o gráfico com título E1 foi utilizado para mostrar a quantidade de canais E1 que estão sendo utilizados simultaneamente durante a replicação. O gráfico com título GSM mostra a quantidade de cartões GSM que as ligações em curso para celular estão consumindo. Já o gráfico VoIP informa a quantidade de ligações DDD utilizando o provedor VoIP que estão em andamento atualmente. Os gráficos E1 e GSM foram configurados para ter o limite inferior de 0 e o limite superior igual ao número máximo de recursos suportados pelos cenários. Como não foram considerados limites de recurso de banda de Internet para as chamadas VoIP, o gráfico VoIP tem limite inferior de 0 e não tem limite superior.

O *label* com o valor "13:00:00" mostra o valor atual do relógio da simulação.

4.8.1.1 – Modelagem da "Duração da Chamada" e "Recursos" no Arena

A Figura 4.6 mostra a tela para configuração da distribuição de probabilidade que representa a duração da chamada e os recursos utilizados durante a chamada. Neste exemplo, a chamada usa 2 canais do tronco E1 e os libera após a duração da chamada. O tempo da chamada é uma variável aleatória gerada pela distribuição LogNormal com parâmetros 123 (*LogMean*) e 155 (*LogStd*).

Figura 4.6 - Tela de configuração do tempo de processamento e recursos.

4.8.1.2 – Modelagem do "Tempo entre Chegada de Chamadas" no Arena

A Figura 4.7 apresenta a tela para configuração do tempo entre chegada das chamadas DDD via VoIP usando módulo *create* do Arena.

Figura 4.7 - Tela de configuração do tempo entre chegada de chamadas.

O campo *Name* é o identificador único do módulo e o valor desse campo aparece no fluxograma do modelo. O campo *Entity Type* contém o nome da entidade que será gerada por esse módulo. O valor desse campo pode ser utilizado em um módulo de decisão.

O campo *Type* define o tipo de chegada que deve ser gerada e os possíveis valores para este campo são: *Random* (usa a distribuição exponencial com a média especificada pelo usuário), *Schedule* (usa a distribuição exponencial com a média especificada pelo módulo *Schedule*), *Constant* (o usuário especifica o valor da constante) ou *Expression* (para usar

outras distribuições de probabilidade). Neste caso, utilizou-se o tipo *Schedule*, pois a chegada das chamadas segue o processo de Poisson e as taxas são diferentes nos diversos períodos do dia. Quando a opção *Schedule* é utilizada, deve-se colocar o nome do *Schedule* que contém as taxas de chegadas no campo *Schedule Name*.

O campo *Entities Arrival* indica o número de entidades que entrarão no sistema por cada chegada. Já o campo *Max Arrival* especifica a quantidade máxima de entidades que podem ser geradas por esse módulo. Quando esse valor é atingido, não há criação de novas entidades desse tipo.

A Figura 4.8 apresenta a ferramenta gráfica utilizada para editar e visualizar as taxas de chegadas que são utilizadas pelo módulo *Schedule* no software Arena.

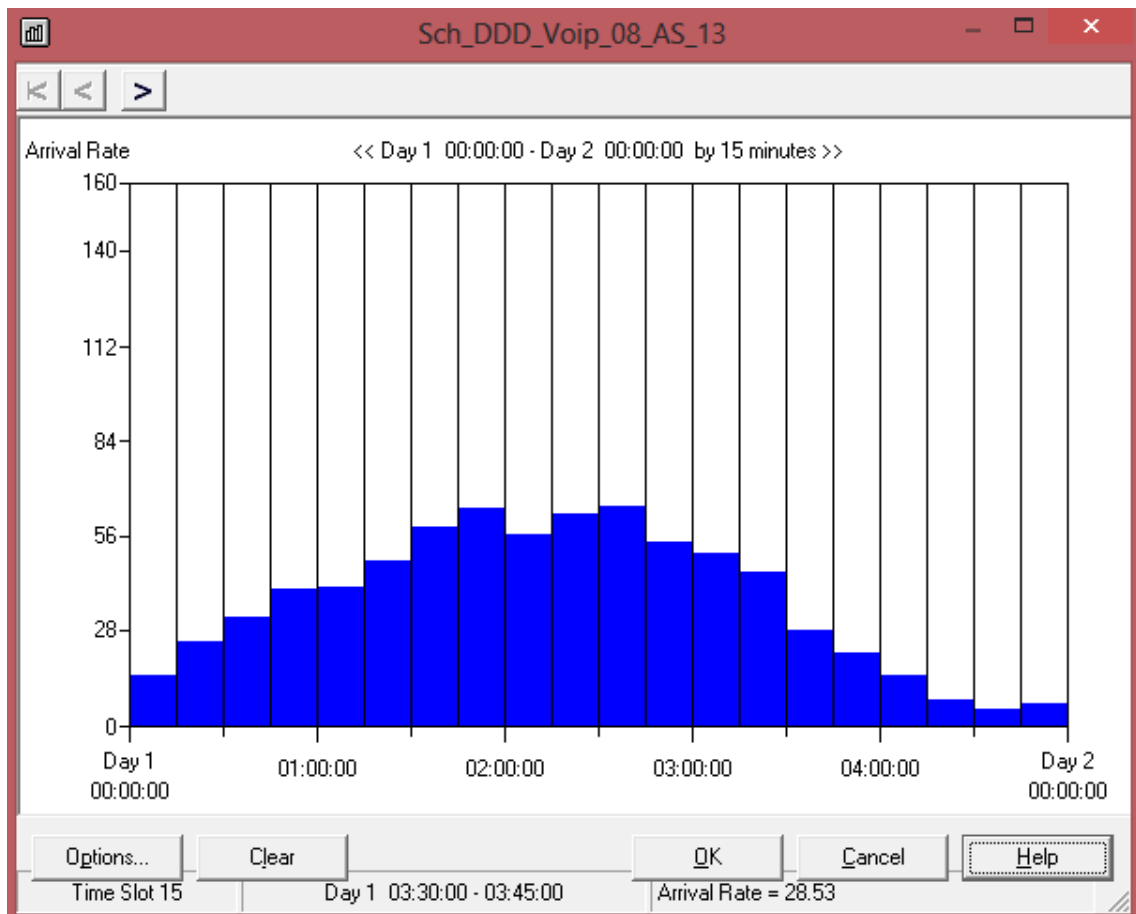


Figura 4.8 - Editor gráfico de *Schedule* do software Arena.

Neste caso, a ferramenta gráfica apresenta as informações das taxas de chegadas das chamadas do tipo I (VoIP) no período da manhã (08:00 às 13:00). O nome do *Schedule* é Sch_DDD_VoIP_08_AS_13 e foram utilizados *slots* de 15 minutos. O valor utilizado em cada *slot* corresponde à quantidade de chamadas por hora naquele período de 15 minutos. As taxas médias de chegadas de chamadas dos 30 dias com maior volume de ligações, agrupadas

em períodos de 15 minutos (ver Tabela 4.4), foram utilizadas para criar os *Schedules*. É importante notar que os valores da Tabela 4.4 precisaram ser multiplicados por 4, pois cada *slot* do *Schedule* deve conter a quantidade de chamadas por hora.

A Figura 4.9 mostra um exemplo do resultado da simulação apresentado pelo software Arena. O resultado indica que foram executadas 1000 replicações. Para cada variável de resposta da simulação (coluna *identifier*) foi apresentado o resultado médio, máximo e mínimo das 1000 replicações. Também é apresentado o valor do semi-intervalo de confiança (*half-width*) para a média. O tempo total de simulação do exemplo foi de 0,58 minutos.

Output Summary for 1000 Replications					
Project: Simulacao do Sistema de Telefonia da UFU			Run execution date : 6/25/2014		
Analyst: Muriel Alves			Model revision date: 3/19/2014		
OUTPUTS					
Identifier	Average	Half-width	Minimum	Maximum	# Replications
E1 AVG Busy	7.4290	.02444	6.2555	8.7081	1000
Canais Voip Max Busy	6.7450	.06504	4.0000	12.000	1000
Chamadas Perdidas DDD	.00000	.00000	.00000	.00000	1000
E1 Max Busy	18.622	.09234	15.000	25.000	1000
GSM AVG Busy	6.0445	.02096	4.9643	7.2086	1000
Canais Voip AVG Busy	1.3487	.01213	.87479	2.1905	1000
Chamadas Perdidas Celular	.65200	.07638	.00000	8.0000	1000
GSM Max Busy	15.395	.05087	13.000	16.000	1000
Ligacoes_Celular_Santa_Monica.NumberIn	817.72	1.7999	712.00	909.00	1000
Ligacoes_Celular_Santa_Monica.NumberOut	814.59	1.7946	708.00	903.00	1000
Ligacoes_DDD_Custo.NumberIn	321.44	1.1086	273.00	375.00	1000
Ligacoes_DDD_Custo.NumberOut	321.44	1.1085	273.00	375.00	1000
Ligacoes_DDD_Voip.NumberIn	203.76	.88595	149.00	259.00	1000
Ligacoes_DDD_Voip.NumberOut	203.39	.88754	149.00	259.00	1000
Canal VoIP.NumberSeized	203.76	.88595	149.00	259.00	1000
Canal VoIP.ScheduledUtilization	.02248	2.0224E-04	.01458	.03651	1000
Canal E1.NumberSeized	1342.2	2.2709	1211.0	1452.0	1000
Canal E1.ScheduledUtilization	.24763	8.1480E-04	.20852	.29027	1000
Placa GSM.NumberSeized	817.07	1.7853	712.00	902.00	1000
Placa GSM.ScheduledUtilization	.37778	.00131	.31027	.45054	1000
System.NumberOut	1339.4	2.2818	1209.0	1450.0	1000
Simulation run time: 0.58 minutes.					
Simulation run complete.					

Figura 4.9 - Exemplo do resultado de uma simulação apresentado pelo Arena.

4.8.2 – Input Analyser

O Input Analyser é uma ferramenta padrão do ambiente Arena que é utilizada para determinar a qualidade da aderência das distribuições de probabilidade aos dados de entrada. Ela suporta os testes de aderência K-S e χ^2 . Além disso, ela fornece os parâmetros estimados da distribuição teórica avaliada no teste.

A Figura 4.10 mostra um exemplo de teste feito com o Input Analyser. Neste caso, o teste de aderência K-S aceitou a hipótese de que os dados seguem a distribuição exponencial

($p\text{-value}=0,0622$). O Input Analyser gera a expressão para ser utilizada no modelo construído no Arena ($0,099 + \text{expo}(116)$).

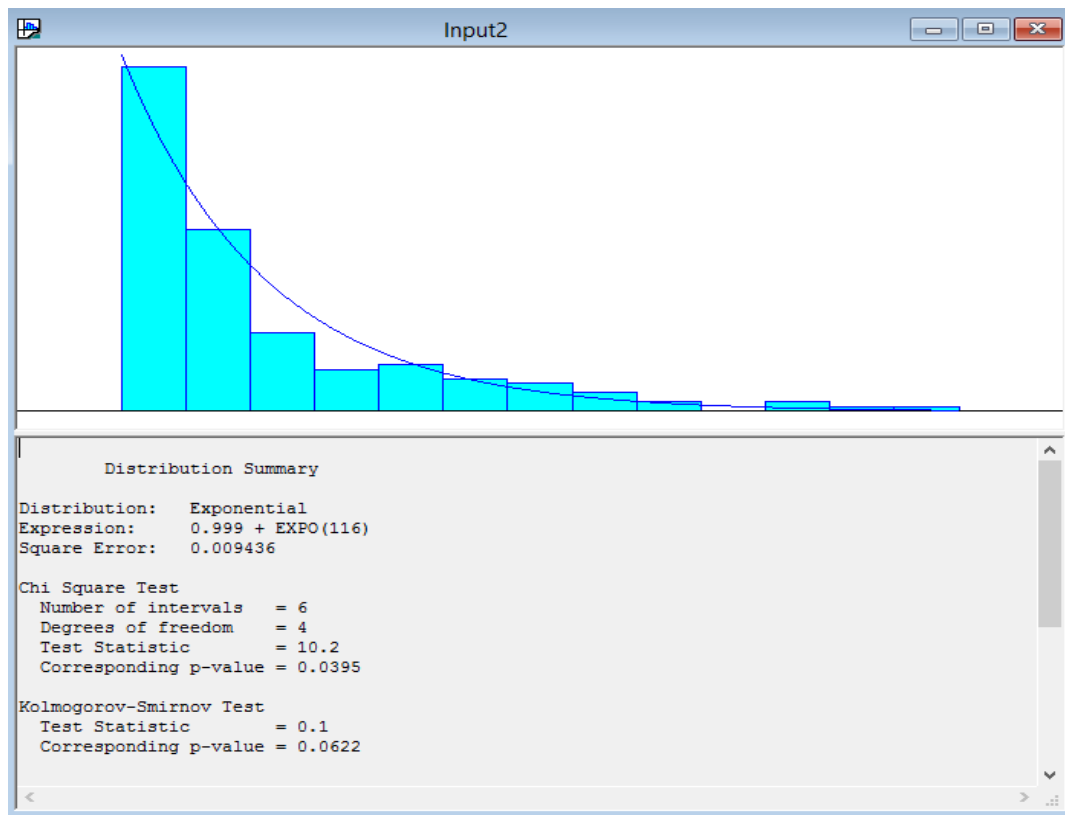


Figura 4.10 - Teste de aderência utilizando o Input Analyser.

4.8.3 – Minitab

O software Minitab [Minitab 2014] versão 14 foi utilizado para fazer o teste de aderência A-D e para gerar os gráficos de probabilidade ($p\text{-}p$ plots).

4.8.4 – Microsoft Excel

O software Microsoft Excel foi utilizado para organizar os dados em planilhas e gerar os gráficos necessários para análise das chamadas.

4.9 – Validação do Modelo

Para validar o modelo de simulação proposto neste estudo, foi necessário verificar se ele representa o sistema de forma precisa [Law 2009]. Primeiramente, foram feitas reuniões com os analistas responsáveis pelo sistema de telefonia da universidade, a fim de apresentar os detalhes do modelo de simulação proposto. O objetivo principal das reuniões foi colher o *feedback* dos analistas para melhorar o modelo.

Além das entrevistas com os analistas, também foi realizada a validação quantitativa do modelo, comparando os resultados das variáveis de interesse obtidos pela simulação com os resultados das mesmas variáveis observados no sistema real.

Para validar quantitativamente o modelo, inicialmente, ele foi simulado utilizando os recursos e tipos de chamadas dos cenários #1 e #2 (ver Tabela 4.9) para obtenção das variáveis de resposta de interesse (ver Tabela 4.8). Depois, foram calculados os valores dessas mesmas variáveis de resposta usando dados reais obtidos através dos CDRs do PABX e do *gateway* de voz. Finalmente, foi realizada a comparação dos valores observados e valores simulados. A qualidade do modelo de simulação proposto foi avaliada com base na diferença dos valores observados com os valores simulados.

Como mencionado na Seção 4.3 (Coleta de Dados), foram utilizados os dados dos 30 dias com maior número de chamadas do ano de 2012 para servir de entrada para o modelo de simulação. Assim, os CDRs desses 30 dias também foram utilizados no processo de validação. A Figura 4.11 apresenta a visão geral do processo de validação do modelo de simulação.

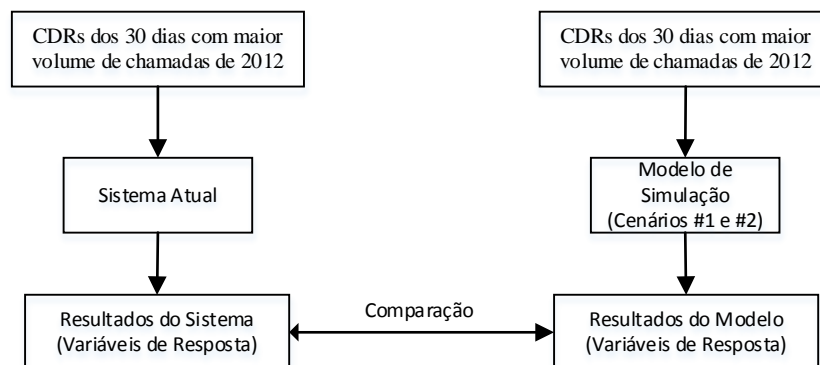


Figura 4.11 - Visão geral do processo de validação do modelo de simulação.

As seguintes verificações foram realizadas no processo de validação do modelo:

- Análise dos valores das variáveis de resposta *VoIP_total*, *PSTN_total* e *Mobile_total* para verificar se as quantidades de chamadas geradas pelo modelo de simulação são similares às quantidades de chamadas observadas no sistema real.
- Análise dos valores das variáveis de resposta *E1_max* e *GSM_max* para verificar se o número máximo de canais E1 e cartões GSM utilizados no sistema real são similares aos valores produzidos pelo modelo de simulação.
- Verificação da similaridade dos valores simulados e observados para a média do número de canais E1 utilizados simultaneamente (*E1_avg*) e a média do número de

cartões GSM utilizados simultaneamente (*GSM_avg*). A Equação 12 foi utilizada para calcular essas duas variáveis,

$$Avg_usage = T^{-1} \times \int_0^T B(t) dt, \quad (12)$$

onde T é o período de tempo analisado (em segundos), e $B(t)$ é a quantidade do recurso alocado (E1 ou GSM) no tempo t .

Os resultados da validação utilizando os cenários #1 e #2 são apresentados na próxima seção.

4.9.1 – Validação do Modelo de Simulação Baseada nos Cenários #1 e #2

O cenário #1 representa o cenário atual do sistema de telefonia da universidade (demanda corrente de chamadas e número corrente de recursos). As chamadas dos tipos I e II foram simuladas e a interconexão do PABX com o *gateway* de voz é realizada através de um tronco E1 (30 canais). A Tabela 4.10 apresenta os valores das variáveis de resposta obtidos através da simulação (coluna “Simulado”) e os valores obtidos através dos dados reais (coluna “Observado”). A coluna “IC” apresenta os limites superiores e inferiores dos intervalos de confiança, ao nível de 95%.

Tabela 4.10 - Resultados simulados e observados do cenário #1.

Variável de Resposta	Manhã			Tarde		
	Simulado	IC	Observado	Simulado	IC	Observado
<i>E1_max</i>	20,51	[20,36 – 20,66]	20,56	19,45	[19,31 - 19,59]	19,60
<i>E1_avg</i>	5,36	[5,34 – 5,38]	5,57	5,46	[5,44 – 5,48]	5,64
<i>E1_rate</i>	17,88%	[17,00 – 18,76]	18,56%	18,22%	[17,34 – 19,1]	18,80%
<i>VoIP_total</i>	189,40	[188,53 – 190,27]	188,9	198,14	[197,27 - 199,01]	197,76
<i>PSTN_total</i>	302,25	[301,14 – 303,36]	301,7	306,20	[305,12 – 307,28]	305,66
<i>LDC_lost</i>	0	[0 – 0]	0	0	[0 – 0]	0

Como explicado na Seção 4.3, a simulação é executada para dois turnos (manhã e tarde), sendo estes os períodos de maior representatividade de uso do sistema de telefonia e *gateway* de voz simulado. Comparando os valores simulados com os valores observados, verifica-se que os valores observados estão dentro do intervalo de confiança para a maioria das variáveis, indicando que o modelo de simulação representa adequadamente o sistema real. As únicas variáveis de resposta que não ficaram dentro do intervalo de confiança foram *E1_avg* (nos períodos da manhã e tarde) e *E1_max* no período da tarde. No caso da variável *E1_avg*, os valores observados ficaram 0,19 (manhã) e 0,16 (tarde) acima do intervalo de confiança. No caso da variável *E1_max*, o valor observado ficou 0,01 acima do intervalo de confiança no período da tarde.

No segundo cenário foram simuladas as chamadas dos tipos I, II e III. Foram utilizados recursos ilimitados de canais E1 e cartões GSM. Neste cenário, foi possível analisar o comportamento do sistema quando atendendo a chamadas para telefones celulares. A Tabela 4.11 apresenta os resultados da simulação e os resultados observados.

Tabela 4.11 - Resultados simulados e observados do cenário #2.

Variável de Resposta	Manhã			Tarde		
	Simulado	IC	Observado	Simulado	IC	Observado
<i>EI_max</i>	29,20	[29,03 – 29,37]	29,43	27,82	[27,67 – 27,97]	27,33
<i>EI_avg</i>	10,51	[10,48 – 10,54]	10,41	10,94	[10,91 – 10,97]	10,84
<i>EI_rate</i>	N/A	N/A	N/A	N/A	N/A	N/A
<i>VoIP_total</i>	189,12	[188,26 – 189,98]	188,9	197,74	[196,87 – 198,61]	197,76
<i>PSTN_total</i>	301,73	[300,67 – 302,79]	301,70	305,97	[304,91 – 307,03]	305,66
<i>LDC_lost</i>	0	[0 – 0]	0	0	[0 – 0]	0
<i>GSM_max</i>	15,22	[15,12 – 15,32]	14,33	14,92	[14,83 – 18,01]	14,10
<i>GSM_avg</i>	5,18	[5,17 – 5,19]	4,84	5,50	[5,49 – 5,51]	5,19
<i>GSM_rate</i>	N/A	N/A	N/A	N/A	N/A	N/A
<i>Mobile_total</i>	700,22	[698,62 – 701,82]	700,06	742,69	[741,01 – 744,37]	742,80
<i>Mobile_lost</i>	0	[0 – 0]	0	0	[0 – 0]	0

Os valores observados da maioria das variáveis de interesse ficaram dentro do intervalo de confiança. Somente os valores observados das variáveis *EI_max*, *EI_avg*, *GSM_max* e *GSM_avg* não ficaram dentro do intervalo de confiança. No entanto, os valores excedidos em relação ao intervalo de confiança foram pequenos. Para a variável *EI_max*, o valor no período da manhã foi de 0,06 acima do intervalo de confiança. No período da tarde o valor foi de 0,34 abaixo do intervalo de confiança. Para a variável *EI_avg* os valores foram 0,07 abaixo do intervalo de confiança nos períodos da manhã e tarde. Para a variável *GSM_max* os valores foram de 0,79 (manhã) e 0,73 (tarde) abaixo do intervalo de confiança. Para a variável *GSM_avg* os valores foram 0,33 e 0,30 abaixo do intervalo de confiança nos períodos da manhã e tarde, respectivamente.

4.10 – Considerações Finais

Neste capítulo foram apresentados os métodos e materiais utilizados para modelar e simular a capacidade do sistema de telefonia apresentado no Capítulo 3.

O processo de validação do modelo contou com o auxílio dos analistas responsáveis pelo sistema de telefonia da universidade. Nesta etapa, os recursos de animação do software Arena foram de extrema importância no processo de interação com os analistas. No próximo capítulo será apresentada uma proposta de melhoria do sistema que foi descoberta durante a visualização da simulação do modelo.

A modelagem do processo de chegada de chamadas foi realizada utilizando o processo de Poisson não estacionário (NHPP). Por isso, foi necessário agrupar os dados relativos às chamadas em períodos de 15 minutos. O recurso *Schedule* do software Arena permitiu a configuração das taxas média de chegada das chamadas para todos os períodos. Assim, não houve a necessidade de executar as replicações separadamente para cada período.

No próximo capítulo será apresentada a análise dos resultados da simulação para planejamento de capacidade dos recursos do *gateway* de voz investigado.

CAPÍTULO 5 – ANÁLISE DOS RESULTADOS DA SIMULAÇÃO

Este capítulo apresenta a análise dos resultados da simulação dos 6 cenários propostos e descritos no Capítulo 4. A simulação dos cenários avaliados tem por principal objetivo avaliar a capacidade do *gateway* de voz, apresentado no Capítulo 3, mediante diferentes configurações de demanda de chamadas e recursos disponíveis.

5.1 – Resultados do Cenário #1

O objetivo do cenário #1 foi avaliar a capacidade do sistema *gateway* de voz frente à demanda atual de chamadas DDD (chamadas dos tipos I e II). Dado que no sistema real o *gateway* de voz e o sistema PABX são interligados usando um tronco E1 (suportando até 30 chamadas de voz simultaneamente), neste cenário foi considerada a configuração atual de canais E1 do sistema real. Os resultados da simulação de tal cenário fornecem aos analistas os valores atuais de utilização de canais E1 do sistema *gateway* de voz. Por isso, esses resultados foram utilizados no processo de validação do modelo (ver Seção 4.9). A Tabela 5.1 apresenta o resultado da simulação do cenário.

Tabela 5.1 - Resultado da simulação do Cenário #1.

Variável de Resposta	Manhã		Tarde	
	Estimado	$\pm HW_{CI}$	Estimado	$\pm HW_{CI}$
<i>E1_max</i>	20,51	0,15	19,45	0,14
<i>E1_avg</i>	5,36	0,02	5,46	0,02
<i>E1_rate</i>	17,88%	0,88%	18,22%	0,88%
<i>VoIP_total</i>	189,4	0,87	198,14	0,87
<i>PSTN_total</i>	302,25	1,11	306,20	1,08
<i>LDC_lost</i>	0	0	0	0

O número máximo de canais E1 utilizados simultaneamente foi de 20,51 no período da manhã, enquanto no período da tarde foi de 19,45. Esses valores mostram que, nos períodos de picos de chamadas, o sistema opera sem chegar ao seu limite, visto que este suporta até 30 chamadas simultâneas.

A média de canais E1 utilizados simultaneamente foi de 5,36 no período da manhã e 5,46 no período da tarde. A diferença entre a média de utilização de canais (*E1_avg*) e o número máximo de canais utilizados (*E1_max*) é explicada pela variabilidade do número de chamadas dentro do período, isto é, há picos de chamadas durante os períodos que elevam o número de canais utilizados simultaneamente.

Apesar de o número máximo de canais E1 utilizados no período da manhã ter sido superior ao período da tarde, o volume de ligações do tipo I e II é maior no período da tarde.

Dessa forma, conclui-se que a demanda por tal recurso (canais E1) é maior no período da tarde, considerando a configuração atual.

Levando em conta a demanda corrente de chamadas DDD (*VoIP_total* e *PSTN_total*), foi observado que 38,5% (manhã) e 39,2% (tarde) das chamadas DDD são completadas usando o provedor VoIP, implicando em redução de custos, visto que as tarifas da operadora VoIP são menores que as tarifas do provedor PSTN.

A taxa média de utilização de canais E1 é menor que 20% da capacidade total, e o número máximo de canais E1 utilizados simultaneamente é de aproximadamente 21 canais. Portanto, conclui-se que os 30 canais atuais do tronco E1 são suficientes para atender a demanda atual de chamadas dos tipos I e II, pois não houve perda de chamada nos dois períodos analisados (*LDC_lost*).

5.2 – Resultados do Cenário #2

Neste cenário foram simuladas as chamadas dos tipos I, II e III. Foram utilizados recursos ilimitados de canais E1 e cartões GSM. Assim, foi possível analisar o comportamento do sistema quando as chamadas para celulares passam pelo *gateway* de voz. Atualmente, chamadas para telefones celulares não estão implementadas no *gateway* de voz, por isso o objetivo deste cenário é avaliar a capacidade do *gateway* de voz para a futura implementação deste novo serviço. A Tabela 5.2 apresenta o resultado da simulação do cenário.

Tabela 5.2 - Resultado da simulação do Cenário #2.

Variável de Resposta	Manhã		Tarde	
	Estimado	$\pm HW_{CI}$	Estimado	$\pm HW_{CI}$
<i>E1_max</i>	29,2	0,17	27,82	0,15
<i>E1_avg</i>	10,51	0,03	10,94	0,03
<i>E1_rate</i>	N/A	N/A	N/A	N/A
<i>VoIP_total</i>	189,12	0,86	197,74	0,87
<i>PSTN_total</i>	301,73	1,06	305,97	1,06
<i>LDC_lost</i>	0	0	0	0
<i>GSM_max</i>	15,22	0,10	14,92	0,09
<i>GSM_avg</i>	5,18	0,01	5,50	0,01
<i>GSM_rate</i>	N/A	N/A	N/A	N/A
<i>Mobile_total</i>	700,22	1,60	742,69	1,68
<i>Mobile_lost</i>	0	0	0	0

O resultado mostrou que o número máximo de canais E1 utilizados simultaneamente, quando comparados com o Cenário #1, aumentou 42% no período da manhã e 43% no período da tarde.

Neste cenário, foi possível descobrir que são necessários 16 cartões GSM para atender a demanda corrente de chamadas para celular. O número máximo de cartões GSM utilizado simultaneamente durante o período da manhã é superior ao período da tarde. Isto pode ser observado no gráfico da Figura 4.2, onde a maior taxa de chegada para as chamadas do tipo III ocorre entre 10:30 e 11:30.

Apesar de a quantidade máxima de cartões GSM utilizados simultaneamente (*Mobile_total*) no período da manhã ter sido maior, o volume de ligações para celular no período da tarde foi superior em aproximadamente 6%, quando comparado com o período da manhã.

Esse resultado também mostrou que, nos períodos de pico, o número máximo de canais E1 utilizados simultaneamente fica próximo da capacidade atual do sistema no período da manhã (29,2). Já no período da tarde, o número máximo de canais E1 utilizados simultaneamente foi de 27,82.

Com base nos resultados deste cenário, pode-se concluir que 16 cartões GSM são suficientes para atender a demanda atual de chamadas para celular. Além disso, quando as ligações para celular começarem a utilizar o *gateway* de voz, o sistema passará a utilizar praticamente toda sua capacidade de canais E1 durante os períodos de pico.

5.3 – Resultados do Cenário #3

Neste cenário foi simulada a demanda corrente de chamadas dos tipos I (VoIP), II (PSTN) e III (GSM) utilizando um tronco E1 (30 canais) e 8 cartões GSM. Vale lembrar que, atualmente, as chamadas para celular (tipo III) não passam pelo *gateway* de voz. Assim, a simulação deste cenário visa avaliar a utilização dos recursos do *gateway* de voz considerando também este tipo de chamada.

No Cenário #2 foi verificado que são necessários 16 cartões GSM para atender a demanda corrente de chamadas para celulares. No entanto, neste cenário foi avaliada a porcentagem de perda de chamadas caso sejam usados apenas 8 cartões GSM instalados no *gateway* de voz. A justificativa para usar apenas 8 cartões deve-se ao fato de o computador que executa o *gateway* de voz suportar apenas mais duas placas PCI com 4 *slots* para cartões GSM. Caso seja necessário utilizar mais cartões GSM, então será necessário adquirir um novo hardware com maior número de *slots* PCI. A Tabela 5.3 apresenta o resultado da simulação.

Tabela 5.3 - Resultado da simulação do Cenário #3.

Variável de Resposta	Manhã		Tarde	
	Estimado	$\pm HW_{CI}$	Estimado	$\pm HW_{CI}$
<i>El_max</i>	26,85	0,13	25,70	0,13
<i>El_avg</i>	9,94	0,02	10,35	0,02
<i>El_rate</i>	33,15%	0,01%	34,50%	0,01%
<i>VoIP_total</i>	189,24	0,85	197,37	0,83
<i>PSTN_total</i>	301,60	1,09	305,96	1,05
<i>LDC_lost</i>	0,11	0,02	0,05	0,02
<i>GSM_max</i>	8	0	8	0
<i>GSM_avg</i>	4,60	0,01	4,89	0,01
<i>GSM_rate</i>	57,50%	0,01%	61,16%	0,01%
<i>Mobile_total</i>	700,52	1,57	743,81	1,69
<i>Mobile_lost</i>	77,40	1,05	82,80	1,12

A variável *GSM_max* atingiu o valor máximo de utilização em todas as replicações dos períodos da manhã e tarde, o que pode ser comprovado pelo valor 0 do semi-intervalo de confiança ($\pm HW_{CI}$) nos dois períodos.

A variável de resposta *Mobile_lost* aponta uma perda de aproximadamente 77 chamadas no período da manhã e 82 chamadas no período da tarde. Esses valores representam, respectivamente, 11,04% e 11,13% de perda de chamadas para telefones celulares.

Como os percentuais de perda de chamadas para celular foram de aproximadamente 11%, considerando que no sistema atual não há perda de ligações para celular, não se recomenda a implementação de apenas 8 cartões GSM no *gateway* de voz para atender a demanda corrente de chamadas para celular.

5.4 – Resultados do Cenário #4

O objetivo deste cenário é avaliar o comportamento do sistema caso haja os seguintes aumentos no número de chamadas: tipo I (3%), tipo II (5%) e tipo III (10%). Foi considerado um tronco E1 (30 canais) e 16 cartões GSM. A Tabela 5.4 apresenta o resultado da simulação.

O número máximo de canais E1 utilizados simultaneamente foi de 29,14 (manhã) e 28,45 (tarde). Devido ao aumento de chamadas, fazendo a comparação com os resultados do cenário #2, a taxa média de utilização de canais E1 (*El_avg*) aumentou de 10,51 para 11,28 no período da manhã, e de 10,94 para 11,73 no período da tarde. A perda de chamadas para celulares foi de aproximadamente 1,68 (manhã) e 0,86 (tarde). A perda de chamadas DDD foi de 0,88 (manhã) e 0,39 (tarde).

Tabela 5.4 - Resultado da simulação do cenário #4.

Variável de Resposta	Manhã		Tarde	
	Estimado	$\pm HW_{CI}$	Estimado	$\pm HW_{CI}$
<i>El_max</i>	29,14	0,07	28,45	0,09
<i>El_avg</i>	11,28	0,03	11,73	0,03
<i>El_rate</i>	37,60%	0,01%	39,11%	0,01%
<i>VoIP_total</i>	194,38	0,82	203,69	0,88
<i>PSTN_total</i>	317,40	1,09	321,50	1,16
<i>LDC_lost</i>	0,88	0,08	0,39	0,05
<i>GSM_max</i>	15,50	0,04	15,35	0,05
<i>GSM_avg</i>	5,70	0,02	6,03	0,02
<i>GSM_rate</i>	35,62%	0,01%	37,69%	0,01%
<i>Mobile_total</i>	771,10	1,78	816,84	1,79
<i>Mobile_lost</i>	1,68	0,14	0,86	0,09

Portanto, considerando o aumento de demanda de chamadas proposto neste cenário, o *gateway* passaria a trabalhar no limite da sua capacidade durante os períodos de pico de chamadas.

5.5 – Resultados do Cenário #5

Como foi explicado no Capítulo 4, para completar uma chamada do tipo II são necessários 2 canais do tronco E1. Isso ocorre porque este tipo de chamada passa pelo *gateway* de voz e volta para o PABX antes de ser encaminhada para o provedor PSTN.

Neste cenário, foi avaliada uma proposta de otimização que possibilita a utilização de somente 1 canal do tronco E1 para cada chamada do tipo II.

O PABX possui a funcionalidade de encaminhar uma chamada para uma rota secundária, caso a rota primária não seja capaz de completar a chamada. Isso ocorre quando a rede pela qual a chamada foi primeiramente encaminhada não atende ao número chamado. Em [ITU-T 1998], esta situação é denominada "*termination code number 3*".

Baseado nesta funcionalidade, a otimização aqui proposta realiza a programação do *gateway* de voz para retornar o "*termination code number 3*" para o PABX quando o *gateway* não conseguir completar a chamada usando o provedor VoIP, evitando a necessidade de rotear a chamada de volta para o PABX. Assim, as chamadas do tipo II usam apenas 1 canal E1 por dois segundos, tempo necessário para o *gateway* consultar a tabela de rotas e finalizar a chamada usando o "*termination code number 3*". Ao receber o código de finalização da chamada, o PABX faz uma nova discagem para o número de destino usando o provedor PSTN.

Neste cenário, foi avaliado o comportamento do sistema considerando o procedimento de terminar as chamadas do tipo II descrito acima. Foram simuladas as chamadas do tipo I, II

e III com o aumento de demanda considerada no Cenário #4, utilizando um tronco E1 (30 canais) e 16 cartões GSM. A Tabela 5.5 apresenta os resultados da simulação deste cenário.

Tabela 5.5 - Resultado da simulação do Cenário #5.

Variável de Resposta	Manhã		Tarde	
	Estimado	$\pm HW_{CI}$	Estimado	$\pm HW_{CI}$
<i>E1_max</i>	19,09	0,09	18,62	0,09
<i>E1_avg</i>	7,00	0,02	7,42	0,02
<i>E1_rate</i>	23,36%	0,07%	24,76%	0,08%
<i>VoIP_total</i>	194,48	0,85	203,76	0,88
<i>PSTN_total</i>	317,38	1,08	321,44	1,10
<i>LDC_lost</i>	0	0	0	0
<i>GSM_max</i>	15,50	0,05	15,39	0,05
<i>GSM_avg</i>	5,69	0,01	6,04	0,02
<i>GSM_rate</i>	35,59%	0,01%	37,77%	0,01%
<i>Mobile_total</i>	770,58	1,77	817,72	1,79
<i>Mobile_lost</i>	1,02	0,10	0,65	0,07

Considerando a otimização proposta, observa-se que a taxa de utilização de canais E1 foi de 23% no período da manhã e 24% no período da tarde. No Cenário #4, o qual não considera a otimização, a taxa de utilização foi de 37% no período da manhã e 39% no período da tarde.

Portanto, a proposta de otimização reduziu a taxa de utilização de canais E1 em mais de 14% quando comparado com a configuração atual. Além disso, a otimização permitiu que o *gateway* de voz suportasse o incremento de demanda para as chamadas dos tipos I e II, o que não ocorre no cenário #4.

A vantagem da otimização também pode ser verificada através do número máximo de canais E1 utilizados simultaneamente (*E1_max*). No cenário #4, que não considera a otimização, o número máximo de canais E1 utilizados simultaneamente foi de 29,14 (manhã) e 28,45 (tarde), ficando próximo do limite de 30 canais e com perda de chamadas DDD nos períodos da manhã e tarde. Já no cenário #5, que considera a otimização, o número máximo de canais E1 utilizados simultaneamente foi de 19,09 (manhã) e 18,62 (tarde) e não houve perda de chamadas DDD.

Na Seção 5.6 serão apresentados os resultados da simulação de um cenário que possibilitou quantificar o número de novas chamadas suportadas pelo *gateway* quando a otimização proposta é utilizada.

5.6 – Resultados do Cenário #6

O objetivo do cenário #6 é quantificar o que o valor percentual de 14% a mais de disponibilidade dos canais E1, oferecido pela otimização apresentada no cenário #5,

representa em número de chamadas a mais que podem ser completadas utilizando o *gateway* de voz.

Para isso, foram simuladas as chamadas dos tipos I, II e III utilizando um tronco E1 (30 canais) e cartões GSM ilimitados. As taxas médias de chegadas dos três tipos de chamadas foram ampliadas para que elas ocupassem os canais E1 durante todo o tempo, fazendo com que a taxa de utilização de canais E1 chegasse em 100%.

A Tabela 5.6 apresenta os resultados da simulação considerando o modelo sem otimização e o modelo com a proposta de otimização implementada.

Tabela 5.6 - Resultado da simulação do Cenário #6.

Variável de Resposta	Modelo sem Otimização				Modelo com Otimização			
	Manhã		Tarde		Manhã		Tarde	
	Estimado	HW _{CI}	Estimado	HW _{CI}	Estimado	HW _{CI}	Estimado	HW _{CI}
<i>E1_avg</i>	29,94	2,3E-04	29,95	1,9E-04	29,94	2,3E-04	29,94	1,9E-04
<i>E1_rate</i>	99,81%	7,7E-06	99,83%	6,3E-06	99,80%	7,8E-06	99,82%	6,6E-06
<i>VoIP_total</i>	845,6	4,70	853,89	4,63	864,98	4,95	869,92	4,79
<i>PSTN_total</i>	66,42	1,75	62,03	1,86	1441,6	6,56	1408,1	7,62
<i>Mobile_total</i>	3445,5	11,26	3441,2	11,29	3531,6	10,87	3524,0	10,04

Note que tanto no modelo sem otimização quanto no modelo com otimização, a taxa de utilização de canais E1 foi maior que 99,8% nos períodos da manhã e da tarde. Isso indica que nos dois casos o *gateway* de voz estava utilizando praticamente sua capacidade máxima de canais E1 durante todo o período considerado na simulação.

O número de chamadas diárias atendidas do modelo sem otimização foi de aproximadamente 8714. Já no cenário com otimização esse número foi aproximadamente 11640. Portanto, a otimização possibilita que o *gateway* de voz atenda 2926 novas chamadas diárias sem adição de novos recursos.

Além disso, pode-se observar que o número de chamadas DDD completadas pelo provedor PSTN no modelo sem otimização (que exige dois canais do tronco E1 para serem atendidas) é bem mais baixo em relação ao modelo otimizado. Isso ocorre porque quando o *gateway* está trabalhando com carga máxima, dificilmente ele terá dois canais do seu tronco E1 disponíveis para completar essas chamadas.

5.7 – Considerações Finais

Neste capítulo foram apresentadas as análises dos resultados das simulações para verificar o comportamento da utilização dos recursos (canais E1 e cartões GSM) do *gateway*

de voz investigado nesta pesquisa. Foram simulados 6 cenários, variando a quantidade de recursos utilizados em cada um deles.

Os cenários #1, #2 e #3 foram simulados utilizando a demanda corrente de chamadas. Já os cenários #4 e #5 foram simulados considerando um aumento da demanda de chamadas com o objetivo de verificar o comportamento do *gateway* caso exista incremento do volume de ligações. O cenário #6 foi simulado para quantificar o número de novas chamadas que podem ser completadas pelo *gateway* de voz caso seja implementada a proposta de otimização discutida no cenário #5, visando reduzir a taxa de ocupação de canais E1.

Uma das vantagens da simulação, citada em [Law e Kelton 1991], é a possibilidade de avaliar uma alteração no modelo do sistema antes de concretizar a alteração no sistema real. Neste estudo, a simulação permitiu avaliar a quantidade de cartões GSM que são necessários instalar no *gateway* de voz para que ele atenda à nova demanda de chamadas para celular antes de fazer qualquer tipo de investimento. Além disso, foi possível avaliar uma proposta de alteração de configuração de rotas no *gateway* de voz, implementada no modelo de simulação, para avaliar a melhoria que esta nova configuração pode trazer caso seja implementada no sistema real.

CAPÍTULO 6 – CONCLUSÕES DA PESQUISA

6.1 – Principais Resultados

Neste trabalho, foi apresentada uma abordagem baseada em simulação computacional, para avaliar e planejar a capacidade de um *gateway* de voz. Tal abordagem foi aplicada em um estudo de caso real, o qual considerou um sistema *gateway* de voz em operação que atende uma Universidade Federal composta de seis campi.

Devido ao elevado número de detalhes envolvidos neste tipo de trabalho, o uso de abordagens analíticas tornaria complexa a obtenção de respostas detalhadas para o planejamento de capacidade dos recursos de sistemas *gateway* de voz. Por exemplo, o modelo analítico Erlang C, bastante utilizado no setor de telefonia, faz a suposição de que a variável aleatória que representa o tempo de serviço (duração das chamadas) seja proveniente de uma distribuição exponencial. No estudo de caso realizado nesta pesquisa, verificou-se que as variáveis aleatórias que representam os tempos de serviço seguem uma distribuição LogNormal para os três tipos de chamadas consideradas. Portanto, para usar o modelo Erlang-C seria necessário aproximar o modelo aderente aos dados por meio da distribuição exponencial, ou equivalente, o que nem sempre é possível.

Alternativamente, o modelo de simulação proposto permitiu incorporar todos os detalhes de um sistema real necessários para analisar diferentes cenários de demanda. Neste estudo, foram avaliados 6 cenários de interesse. Os recursos de visualização e animação do modelo de simulação também foram de grande importância durante as entrevistas com os engenheiros na etapa de validação do modelo proposto. Esses recursos possibilitaram uma melhor interação com os engenheiros durante a etapa de eliciação de conhecimento para o entendimento dos diversos detalhes a cerca da modelagem do sistema investigado. Ademais, a visualização e animação do modelo facilitou identificar a otimização do sistema proposta no cenário #5, a qual propiciou uma redução na taxa de utilização do tronco E1 em 14%. Essa redução representa um acréscimo de 2926 novas chamadas diárias sem novos investimentos;

A análise dos resultados dos 6 cenários propostos permitiu atender a todos os objetivos específicos elencados no Capítulo 1. Além disso, a abordagem adotada permite a rápida adaptação do modelo de simulação para responder novas questões relacionadas, tais como:

- Demanda máxima suportada pela configuração atual de recursos;
- Banda de Internet necessária caso haja incremento de chamadas utilizando o provedor VoIP;

- Quantidade de canais E1 necessários para completar todas as chamadas DDD utilizando o provedor PSTN (em caso de falha do provedor VoIP).

Com o estudo de caso desenvolvido neste trabalho, foi possível observar na prática as vantagens da simulação computacional em possibilitar a análise flexível de alterações no modelo do sistema antes de concretizá-las no sistema real. Por exemplo, neste trabalho a simulação permitiu avaliar a quantidade de cartões GSM que são necessários para atender a uma nova demanda de chamadas para celular antes de fazer qualquer tipo de investimento. Além disso, foi possível avaliar uma proposta de alteração da configuração atual das rotas no *gateway* de voz, testada no modelo de simulação, para medir a melhoria dessa mudança no sistema real.

6.2 – Limitações da Pesquisa

Durante a coleta de dados, deparou-se com o problema de falta de informações sobre as chamadas do tipo III (GSM) que não foram completadas. O sistema PABX estudado não estava configurado para registrar essas chamadas nos CDRs. As informações das chamadas não completadas são importantes para obter um resultado mais preciso sobre a utilização de recursos, visto que durante uma chamada não completada existe utilização de um canal E1 e um cartão GSM. Apesar de ser um período curto de ocupação destes dois recursos, mesmo assim existe influência nos resultados. Desta forma, como as chamadas não completadas do tipo III (GSM) não foram consideradas no modelo, os resultados dos cenários onde este tipo de chamada foi simulado (#2, #3, #4, #5 e #6) forneceram valores que não consideraram a utilização de cartões GSM e canais E1 durante o período (*ringing*) das chamadas não completadas.

Outra limitação da pesquisa é o fato de o estudo de caso ter sido realizado com apenas um tipo de sistema *gateway* de voz. A utilização dos métodos e a forma como foi realizada a modelagem e simulação durante esta pesquisa poderão servir de base para outros pesquisadores que desejam modelar e simular outros sistemas de telefonia.

6.3 – Dificuldades Encontradas

Como o tamanho das amostras das chamadas eram grandes (maiores que 11.000 registros CDR), a maior dificuldade encontrada no início da pesquisa foi descobrir distribuições de probabilidade que representassem adequadamente os dados das amostras utilizando os testes de aderências A-D, K-S e χ^2 . Para contornar essa dificuldade, foi utilizada

de forma complementar a análise gráfica baseada nos gráficos de probabilidade (*p-p plots*) (ver Seção 4.5).

6.4 – Contribuição para a Literatura

Os resultados deste trabalho foram a base para a escrita do artigo intitulado "*Modeling and Simulation Applied to Capacity Planning of Voice Gateways: A Case Study*", o qual foi aceito para publicação nos anais da conferência *Winter Simulation Conference* (Qualis A1 CAPES). O artigo será apresentado em dezembro de 2014 e estará disponível nos anais da conferência a partir dessa data.

6.5 – Trabalhos Futuros

Como trabalho futuro pretende-se melhorar o modelo de simulação proposto, adicionando novas variáveis de reposta, em especial variáveis relacionadas com custos financeiros, tais como economia em custos de chamadas para novos códigos DDD atendidos pelo provedor VoIP e em casos de chamadas DDD para novos campi incorporados ao sistema.

Além disso, para calcular o incremento de demanda, pretende-se utilizar modelos de predição baseados em séries temporais (ex. ARIMA, Holt-Winters, *Exponential Smoothing*).

Objetivando melhorar o modelo de simulação proposto, pretende-se também incluir a simulação da rediscagem por parte do usuário. Essa situação acontece quando um usuário tenta fazer uma chamada e esta não é completada por falta de recursos, o que não foi considerado na proposta apresentada neste trabalho.

REFERÊNCIAS BIBLIOGRÁFICAS

- [Aldor-Noiman *et al.* 2009] Aldor-Noiman, S., Feigin, P. D., e Mandelbaum, A. (2009). “Workload Forecasting for a Call Center: Methodology and a Case Study”. *The annals of Applied Statistics*. Vol. 3, no. 4, pp. 1403–1447.
- [Akhtar e Latif 2010] Akhtar, S. e Latif, M. (2010). “Exploiting Simulation for Call Centre Optimization”. In: *Proceedings of the World Congress on Engineering*. pp. 2112–2117.
- [Anderson e Darling 1954] Anderson, T. W. e Darling, D. A. (1954). “A Test of Goodness of Fit”. *Journal of the American Statistical Association*. Vol. 49, no. 268, pp. 765–769.
- [Arena 2014] Arena Simulation Software. <http://www.arenasimulation.com/>. Acessado em 10/06/2014.
- [Avramidis *et al.* 2004] Avramidis, A. N., Deslauriers, A., e L’Ecuyer, P. (2004). “Modeling Daily Arrivals to a Telephone Call Center”. *Management Science*. Vol. 50, no. 7, pp. 896–908.
- [Barbetta *et al.* 2010] Barbetta, P. A., Reis, M. M., e Bornia, A. C. (2010). “Estatística para cursos de engenharia e informática”. 3ª ed. Editora Atlas. São Paulo-SP.
- [Boero *et al.* 2005] Boero, G., Smith, J., e Wallis, K. F. (2005). “The Sensitivity of the Chi-Squared Goodness-of-Fit Test to the Partitioning of Data”. *Econometric Reviews*. Vol. 23, no. 4, pp. 341–370.
- [Bouzada 2009] Bouzada, M. A. C. (2009). “Dimensioning a Call Center: Simulation or Queue Theory?”. *The Flagship Research Journal of International Conference of the Production and Operations Management Society*. Vol. 2, no. 2, pp. 34–46.
- [Chambers *et al.* 1983] Chambers, J. M., Cleveland, W., Kleiner, B., e Tukey, P.A. (1983). “Graphical Methods for Data Analysis”. 1ª ed. Editora Wadsworth International Group.
- [Chromy *et al.* 2011] Chromy, E., Misuth, T., e Kavacky, M. (2011). “Erlang C Formula and its use in the Call Centers”. *Informational and Communication Technologies and Services*. Vol. 9, no. 1, pp. 7–13.
- [Costa Neto 1977] Costa Neto, P. L. O. (1977). “Estatística”. 1ª ed. Editora E. Blucher. São Paulo-SP.
- [Digium 2014] Digium. <http://www.digium.com/>. Acessado em 13/05/2014.
- [Franzese *et al.* 2009] Franzese, L. A. G., Fioroni, M. M., Botter, R. C., e Freitas Filho, P. J. (2009). “Comparison of Call Center models”. In: *Proceedings of the 2009 Winter Simulation Conference*. pp. 2963–2970.
- [Freitas 2010] Freitas, M. A. S. (2010). “Fundamentos do gerenciamento de serviços de TI: preparatório para a certificação ITIL V3 Foundation”. 1ª ed. Editora Brasport.
- [Freitas Filho 2008] Freitas Filho, P. J. (2008). “Introdução à Modelagem e Simulação de Sistemas: com Aplicações Arena”. 2ª ed. Editora Visual Books. Florianópolis-SC.
- [Goldsman 2007] Goldsman, D. (2007). “Introduction to simulation”. In: *Proceedings of the 2007 Winter Simulation Conference*. pp. 26–37.
- [Goode 2002] Goode, B. (2002). “Voice Over Internet Protocol (VoIP)”. In *Proceedings of the IEEE*. Vol. 90, no. 9, pp. 1495–1517.

- [Green *et al.* 2003] Green, L. V., Kolesar, P. J., e Soares, J. (2003). “An improved heuristic for staffing telephone call centers with limited operating hours”. *Production and Operations Management*. Vol. 12, no. 1, pp. 46-61.
- [Ibrahim e L’Ecuyer 2013] Ibrahim, R. e L’Ecuyer, P. (2013). “Forecasting Call Center Arrivals: Fixed-Effects, Mixed-Effects, and Bivariate Models”. *Manufacturing & Service Operations Management*. Vol. 15, no. 1, pp. 72–85.
- [Ibrahim *et al.* 2012] Ibrahim, R., L’Ecuyer, P., Regnard, N., e Shen, H. (2012). “On the modeling and forecasting of call center arrivals”. In: *Proceedings of the 2012 Winter Simulation Conference*. pp. 1–12.
- [ITU-T 1998] "Usage of Cause and Location in the Digital Subscriber Signaling System No. 1 and the Signaling System No. 7 ISDN User Part." ITU-T Q.850.
- [Jongloed e Koole 2001] Jongbloed, G. e Koole, G. (2001). “Managing uncertainty in call centers using Poisson mixtures”. *Applied Stochastic Models in Business and Industry*. Vol. 17, no. 4, pp. 307–318.
- [Kelton *et al.* 2007] Kelton, W. D., Sadowski, R. P., e Sturrock, D. T. (2007) “Simulation With Arena”. 4^a ed. Editora McGraw-Hill.
- [Law 2009] Law, A. M. (2009). “How to build valid and credible simulation models”. In: *Proceedings of the 2009 Winter Simulation Conference*. pp. 24–33.
- [Law e Kelton 1991] Law, A. M. e Kelton, W. D. (1991). “Simulation Modeling & Analysis”. 2^a ed. Editora McGraw-Hill.
- [Lehmer 1951] Lehmer, D. H. (1951). “Mathematical Methods in Large-Scale Computing Units”. *Annals of the Computation Laboratory of Harvard University*. Vol. 26, pp. 141-146.
- [Liao *et al.* 2012] Liao, S., Koole, G., Delft, C., e Jouini, O. (2012). “Staffing a call center with uncertain non-stationary arrival rate and flexibility”. *OR Spectrum*. Vol. 34, no. 3, pp. 691–721.
- [Massey 1951] Massey, F. (1951). “The Kolmogorov-Smirnov Test for Goodness of Fit”. *Journal of the American Statistical Association*. Vol. 46, no. 253, pp. 68–78.
- [Mehrotra e Fama 2003] Mehrotra, V. e Fama, J. (2003). “Call Center simulation modeling: methods, challenges, and opportunities”. In: *Proceedings of the 2003 Winter Simulation Conference*. pp. 135–143.
- [Minitab 2014] Minitab. <http://www.minitab.com/>. Acessado em 20/06/2014.
- [NIST/SEMATECH 2014] NIST/SEMATECH e-Handbook of Statistical Methods. <http://www.itl.nist.gov/div898/handbook/>. Acessado em 25/06/2014.
- [Philips 2014] Philips. <http://www.philips.com/>. Acessado em 10/06/2014.
- [RNP 2014] Rede Nacional de Ensino e Pesquisa. <http://www.rnp.br/>. Acessado em 20/06/2014.
- [Robbins e Harrison 2008] Robbins, T. R. e Harrison, T. P. (2008). “A simulation based scheduling model for call centers with uncertain arrival rates”. In: *Proceedings of the 2008 Winter Simulation Conference*. pp. 2884–2890.
- [Robbins e Medeiros 2010] Robbins, T. R. e Medeiros, D. J. (2010). “Does the Erlang C model fit in real call centers?” In: *Proceedings of the 2010 Winter Simulation Conference*. pp. 2853–2864.

- [Rosenberg *et al.* 2002] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, e M. Schooler, E. (2002). "RFC SIP: Session Initiation Protocol". IETF. Available: <http://www.ietf.org/rfc/rfc3261.txt/>.
- [Sanchez 2007] Sanchez, P. J. (2007). "Fundamentals of simulation modeling". In: Proceedings of the 2007 Winter Simulation Conference. pp. 54-62.
- [Santner *et al.* 2003] Santner, T. J., Williams, B. J e Notz, W. (2003). "The design and analysis of computer experiments". Editora Springer. New York.
- [Schulzrinne *et al.* 2003] Schulzrinne, H., Casner, S., Frederick, R., e Jacobson, V. (2003). "RFC RTP: A transport protocol for real-time application". IETF. Available: <http://www.ietf.org/rfc/rfc3550.txt/>.
- [Sencer e Ozel 2013] Sencer, A. e Ozel B. (2013). "A simulation-based decision support system for workforce management in call centers". Simulation: Transactions of the Society for Modeling and Simulation International. Vol. 89, no. 4, pp. 481-497.
- [Steinmann e Freitas Filho 2013] Steinmann, G. e Freitas Filho, P. J. (2013). "Using simulation to evaluate call forecasting algorithms for Inbound Call Centers". In: Proceedings of the 2013 Winter Simulation Conference. pp. 1132-1139.
- [Stephens 1974] Stephens, M. A. (1974). "EDF Statistics for Goodness of Fit and Some Comparisons". Journal of the American Statistical Association. Vol. 69, no. 347, pp. 730-737.
- [Stephens 1976] Stephens, M. A. (1976). "Asymptotic Results for Goodness-of-Fit Statistics with Unknown Parameters". Annals of Statistics. Vol. 4, no. 2, pp. 357-369.
- [Stephens 1977a] Stephens, M. A. (1977). "Goodness of Fit for the Extreme Value Distribution". Biometrika. Vol. 64, no. 3, pp. 583-588.
- [Stephens 1977b] Stephens, M. A. (1977). "Goodness of Fit with Special Reference to Tests for Exponentiality". Technical Report No. 262. Department of Statistics, Stanford University, Stanford, CA.
- [Stephens 1979] Stephens, M. A. (1979). "Tests of Fit for the Logistic Distribution Based on the Empirical Distribution Function". Biometrika. Vol. 66, no. 3, pp. 591-595.
- [Takakuwa e Okada 2005] Takakuwa, S. e Okada, T. (2005). "Simulation analysis of inbound call center of a city-gas company". In: Proceedings of the 2005 Winter Simulation Conference. pp. 2026-2033.
- [Taylor 2003] Taylor, J. W. (2003). "Short-term electricity demand forecasting using double seasonal exponential smoothing". Journal of The Operational Research Society. Vol. 54, no. 8, pp. 799-805.
- [Taylor 2008] Taylor, J. W. (2008). "A comparison of Univariate Time Series Methods for Forecasting Intraday Arrivals at a Call Center". Management Science, Vol. 54, no. 2, pp. 253-265.