UNIVERSIDADE DO EXTREMO SUL CATARINENSE - UNESC

CURSO DE CIÊNCIA DA COMPUTAÇÃO

EDUARDO DE SOUZA ZARDIN

ANÁLISE DE TRÁFEGO DE DADOS EM REDES LOCAIS: ESTUDO DE CASO NA UNESC

CRICIÚMA

2016

EDUARDO DE SOUZA ZARDIN

ANÁLISE DE TRÁFEGO DE DADOS EM REDES LOCAIS: ESTUDO DE CASO NA UNESC

Trabalho de Conclusão de Curso, apresentado para obtenção do grau de Bacharel em Ciência da Computação da Universidade do Extremo Sul Catarinense, UNESC.

Orientador: Prof. MSc. Rogério Antônio Casagrande
Coorientador: Prof. Dr. Kristian Madeira

CRICIÚMA

2016

EDUARDO DE SOUZA ZARDIN

ANÁLISE DE TRÁFEGO DE DADOS EM REDES LOCAIS: ESTUDO DE CASO NA UNESC

Trabalho de Conclusão de Curso aprovado pela Banca Examinadora para obtenção do Grau de Bacharel, no Curso de Ciência da Computação da Universidade do Extremo Sul Catarinense, UNESC, com Linha de Pesquisa em Redes de computadores

Criciúma, 23 de junho de 2016.

BANCA EXAMINADORA

Prof. MSc. Rogério Antônio Casagrande - (UNESC) - Orientador

Prof. Dr. Kristian Madeira - (UNESC) - Coorientador

Prof. MSc. Paulo João Martins - (UNESC)

Prof. Esp. Valter Blauth Junior - (UNESC)

AGRADECIMENTOS

Agradeço a minha mãe, Ineci, por toda ajuda, compreensão e carinho prestados neste momento em que muitas vezes não estive presente.

Aos meus amigos e colegas do Departamento de Tecnologia de Informação da UNESC, pelos esclarecimentos, ajuda, força e compreensão, que possibilitaram a realização desta pesquisa.

Aos amigos e colegas do Curso de Ciência da Computação em especial aos professores, que sempre muito carinhosos e companheiros contribuíram de forma direta ou indireta na realização e conclusão desta pesquisa.

Ao meu orientador professor Rogério Antônio Casagrande e o coorientador Kristian Madeira, por aceitar este desafio e pelos incentivos.

À UNESC, que disponibilizou os recursos necessários para execução da pesquisa e estudos, em especial a Valéria de Araujo gerente do Departamento de Tecnologia da Informação que disponibilizou acesso ao ambiente de rede.

Enfim, a todos que contribuíram para a execução dessa pesquisa, seja pela ajuda ou por uma simples palavra de conforto.

“A menos que modifiquemos nossa maneira de pensar, não seremos capazes de resolver os problemas causados pela forma como nos acostumamos a ver o mundo”.

Albert Einstein

RESUMO

O uso das redes, tornou-se indispensável em qualquer corporação, sendo o principal objetivo o compartilhamento de informações, recursos e serviços. O controle sobre o que trafega em uma rede é o primeiro passo para manter ela operacional e auxiliar na tomada de decisões sobre políticas de investimentos, contingências, entre outras ações estratégicas à empresa. Para esta pesquisa, foram definidos períodos e tempos de amostragens com o objetivo de analisar o tráfego nas redes existentes na Universidade do Extremo Sul Catarinense. Para tanto, foram utilizadas as técnicas de amostragem estratificada, seguida da aleatória simples. Os dados foram obtidos a partir de uma ferramenta de captura de pacotes, no qual sobre cada amostra foram aplicados métodos estatísticos e testes não paramétricos comparando os resultados obtidos nos períodos analisados. A referida pesquisa apresenta, como estudo de caso, o diagnóstico das redes administrativa e de laboratórios de uma instituição de ensino. Os resultados demonstram que o maior tráfego gerado pela rede administrativa foi interno, concentrado no período matutino e na de laboratórios foi por acessos à Internet, com maior fluxo no período noturno. As análises obtidas demonstram que o propósito dessa pesquisa foi alcançado e consequentemente, auxilia nas decisões que possam ser tomadas para obter-se uma rede com melhor desempenho e vazão de informação.

Palavras-chave: Redes. Análise de tráfego. Tráfego de rede. *Sniffer*. Wireshark.

ABSTRACT

The use of networks, has become indispensable in any corporation, with the main objective to share information, resources and services. Control over what travels on a network is the first step to keep it operating and assist in making decisions about investment policies, contingencies, among other strategic actions to the company. For this research, periods were defined and sampling times in order to analyze traffic on existing networks at Universidade do Extremo Sul Catarinense. Therefore, the stratified sampling techniques were used, followed by simple random. Data were obtained from a packet capture tool, in which for each sample were applied statistical methods and nonparametric tests comparing the results obtained in the periods analyzed. Such research has, as a case study, the diagnosis of administrative networks and laboratories of an educational institution. The results show that the bigger traffic generated by the administrative network was internally, concentrated in the morning and in the laboratories was for access to the Internet with greater flow at night. The obtained analyzes show that the purpose of this study was achieved and consequently assists in decisions that may be taken to obtain a network with better performance and flow of information.

Keywords: Networks. Traffic analysis. Network traffic. Sniffer. Wireshark.

LISTA DE ILUSTRAÇÕES

Figura 1 - Arquitetura TCP/IP 19
Figura 2 - Datagrama do protocolo IP 21
Figura 3 - Formato de pacote do protocolo TCP com seus campos de controle e dados que transporta para aplicação 26
Figura 4 - Formato do pacote UDP 27
Figura 5 - Conexão do tipo TELNET e do tipo SSH 29
Figura 6 - Funcionamento do request e response do HTTP 30
Figura 7 - Transferência de mensagens entre Alice e Bob 32
Figura 8 - Usando POP3 para baixar três mensagens 33
Figura 9 - Arquitetura de um *sniffer* 36
Figura 10 - Interface gráfica do Wireshark 37
Figura 11 - Interface gráfica do Microsoft Message Analyzer 38
Figura 12 - TCPDUMP 38
Figura 13 - Interface gráfica Capsa Packet *Sniffer* 39
Figura 14 - Interface gráfica do NetworkMiner 40
Figura 15 - Processo iterativo das pesquisas 41
Figura 16 - Gráfico de 180 dias gerado pelo The Dude 49
Figura 17 - Tráfego da semana anterior a amostragem rede administrativa 50
Figura 18 - Tráfego da semana anterior a amostragem rede laboratório 50
Figura 19 - Número de usuários rede sem fio 52
Figura 20 - Diagrama da rede XXI-A 53
Figura 21 - Diagrama da rede do bloco S 55
Figura 22 - Tela principal do Wireshark 2.0.3 57
Figura 23 - Área destinada para aplicação de filtros 58

LISTA DE TABELAS

Tabela 1 – Período de amostragem da rede administrativa 51
Tabela 2 - Período de amostragem da rede laboratório 51
Tabela 3 - Alguns filtros utilizados para identificação do tráfego 58
Tabela 4 - Tráfego total da rede administrativa 60
Tabela 5 - Tráfego total da rede laboratório 61
Tabela 6 - Tráfego de Navegação (WWW) bloco S 63
Tabela 7 - Tráfego de navegação HTTPS 63
Tabela 8 - Sites mais acessados no período de coleta 64
Tabela 9 - Resumo do tráfego interno bloco S 65
Tabela 10 - Quantidade de protocolos capturados período matutino 66
Tabela 11 - Quantidade de protocolos capturados período vespertino 66
Tabela 12 - Quantidade de protocolos capturados período noturno 67
Tabela 13 - Tráfego de navegação (WWW) XXI-A 68
Tabela 14 - Tráfego de navegação HTTPS XXI-A 68
Tabela 15 - Sites mais acessados no período de coleta XXI-A 70
Tabela 16 - Tráfego interno XXI-A 70
Tabela 17 - Quantidade de protocolos capturados período matutino 71
Tabela 18 - Quantidade de protocolos capturados período vespertino 72
Tabela 19 - Quantidade de protocolos capturados período noturno 72

LISTA DE ABREVIATURAS E SIGLAS

ARP	Address Resolution Protocol
ASCII	American Standard Code for Information Interchange
AT	Armário de Telecomunicação
AUNESC	Associação UNESC
CRC	Cyclic Redundancy Checkl
DNS	Domain Name System
FTP	File Transfer Protocol
GB	*Gigabyte*
Gbps	Gigabits por segundo
HTTP	HiperText Transfer Protocol
IBM	International Business Machines
ICMP	Internet Control Message Protocol
IGMP	Internet Group Management Protocol
IMAP	Internet Message Access Protocol
IP	Internet Protocol
IPv4	Internet Protocol versão 4
IPv6	Internet Protocol versão 6
LAN	Local Area Network
MAC	Media Access Control
MAC OS	Macintosh Operating System
MAN	Metropolitan Area Network
MB	*Megabyte*
NBNS	Netbios Name Service
Mbps	Megabits por segundo
NBSS	Netbios Session Service
NIC	Network Interface Card
POP	Post Office Protocol
POP3	Post Office Protocol versão 3
RARP	Reverse Address Resolution Protocol
RFC	Request for Comments
SMTP	Simple Mail Transfer Protocol
SPSS	Statistical Package for the Social Sciencies

SSH	Secure Shell
TCP	Transmission Control Protocol
TELNET	Terminal Emulator
TFTP	Trivial File Transfer Protocol
TI	Tecnologia da Informação
TTL	Time To Live
UDP	User Datagram Protocol
UFLA	Universidade Federal de Lavras
UNESC	Universidade do Extremo Sul Catarinense
WAN	Wide Area Network
WWW	World Wide Web
XMPP	eXtensible Messaging and Presence Protocol

SUMÁRIO

1 INTRODUÇÃO 14
1.1 OBJETIVO GERAL 14
1.2 OBJETIVOs ESPECÍFICOS 15
1.3 JUSTIFICATIVA 15
1.4 ESTRUTURA DO TRABALHO 16
2 REDES DE COMPUTADORES 17
2.1 REDES LOCAIS E REDES DE LONGA DISTÂNCIA 17
2.2 ARQUITETURA EM CAMADAS – TCP/IP 19
2.3 PROTOCOLOS DE COMUNICAÇÃO 20
2.3.1 Protocolo IP 21
2.3.2 Protocolo IP versão 4 (IPv4) 22
2.3.3 Protocolo IP versão 6 (IPv6) 23
2.3.4 Protocolo Internet Control Message Protocol (ICMP) 24
2.3.5 Protocolo Address Resolution Protocol (ARP) 24
2.3.6 Protocolo Reverse Address Resolution Protocol (RARP) 25
2.3.7 Protocolo TCP 25
2.3.8 Protocolo UDP 27
2.3.9 Protocolo File Transfer Protocol (FTP) e Trivial File Transfer Protocol (TFTP) 28
2.3.10 Protocolo Terminal Emulator (Telnet) e protocolo Secure SHell (SSH) ..29
2.3.11 Protocolo HiperText Transfer Protocol (HTTP) 29
2.3.12 Correio eletrônico (e-mail) 31
2.3.12.1 Simple Mail Transfer Protocol (SMTP) 31
2.3.12.2 Post Office Protocol, version 3 (POP3) 32
2.3.12.3 Internet Message Access Protocol (IMAP) 34
3 FERRAMENTAS DE MONITORAMENTO 35
3.1 *Sniffer* 35
3.2 Principais *sniffers* e suas funcionalidades 36
3.2.1 Wireshark 37
3.2.2 Microsoft Message Analyzer 37
3.2.3 Tcpdump 38
3.2.4 Capsa Packet *Sniffer* 39

3.2.5 NetworkMiner....39
4 MÉTODOS ESTATÍSTICOS....41
4.1 TESTE U DE MANN-WHITNEY....41
4.2 TESTE H DE KRUSKAL-WALLIS....42
4.3 METODO DE DUNN....43
4.4 TESTE KOLMOGOROV-SMIRNOV....43
5 TRABALHOS CORRELATOS....45
5.1 DIAGNÓSTICO DO TRÁFEGO DE REDE DE LABORATÓRIOS DE INFORMÁTICA. ESTUDO DE CASO: UNIVERSIDADE DO EXTREMO SUL CATARINENSE....45
5.2 ANÁLISE DA REDE SOB O PONTO DE VISTA DO CONTROLE DE INFORMAÇÕES E TRÁFEGO ESTUDO DE CASO: TSA QUÍMICA DO BRASIL....45
5.3 DIAGNÓSTICO DO TRÁFEGO DE REDE WEB E ANÁLISE DA BASE DE DADOS GERADA PELO MICROSOFT INTERNET SECURITY AND ACCELERATION 2006: ESTUDO DE CASO NA SATC....46
5.4 GERENCIAMENTO E MONITORAÇÃO DE REDES DE COMPUTADORES UTILIZANDO-SE ZABBIX....46
6 ANÁLISE NO TRÁFEGO DA REDE ADMINISTRATIVA E DE LABORATÓRIO....47
6.1 METODOLOGIA DE ANÁLISE....47
6.2 DEFINIÇÃO DE HORÁRIOS E TEMPO DE COLETA....47
6.2.1 Amostragem estratificada e aleatória simples....48
6.2.2 Horários e amostragem....48
6.3 CENÁRIO....52
6.3.1 Estrutura da rede....52
6.3.2 Aplicação da ferramenta Wireshark....56
6.3.3 Aplicação de filtros sobre as amostras....57
6.3.4 Armazenamento dos dados....59
6.4 RESULTADOS OBTIDOS....60
7 CONCLUSÃO....74
REFERÊNCIAS....78
APÊNDICE(S)....84
APÊNDICE A – ARTIGO....85
ANEXO(S)....92
ANEXO A – AUTORIZAÇÃO DE CAPTURA DE DADOS....93

1 INTRODUÇÃO

A infraestrutura de redes de computadores está ficando cada vez mais importante para as organizações e trata-se de um serviço indispensável que precisa estar 100% (cem por cento) operacional (LIMA, 2014).

Recursos computacionais nos permitem iniciar um trabalho no escritório, visualizá-lo no smartphone e finalizá-lo em casa no notebook ou no desktop. Com o avanço da tecnologia, é notável a necessidade de transmissão de informações entre diversos dispositivos tais como: desktops, notebooks, tablets e smartphones (COMER, 2006).

Com isso, Kurose e Ross (2006) destacam a necessidade do monitoramento e controle do fluxo de dados pois profissionais acabam desconhecendo o que trafega em sua própria rede. Como consequência, tomadas de decisões a respeito do desempenho de uma rede são definidas de forma errônea.

Levando em consideração essas informações, observa-se a necessidade de profissionais da área da tecnologia obterem um melhor detalhamento do tráfego de dados da rede. Analisar congestionamento de dados, identificando o causador de determinado problema que pode ser um usuário ou algo que está trafegando pela rede naquele momento, assim como os períodos mais críticos de acessos e a causa desses problemas.

Existem ferramentas de monitoramento que auxiliam esses profissionais, gerando informações dos dados obtidos, através de gráficos e outras formas de visualização, e um exemplo destas ferramentas é o *sniffer* que faz a captura de pacotes na rede e possibilita a verificação de seu conteúdo.

Neste trabalho foi feito uma avaliação na rede da Universidade do Extremo Sul Catarinense (UNESC) por meio de ferramentas de monitoramento de rede, com vistas a propor melhorias na infraestrutura.

1.1 OBJETIVO GERAL

Monitorar o tráfego de dados em uma rede local em busca de evidências de uso inadequado de recursos.

1.2 OBJETIVOS ESPECÍFICOS

Os seguintes objetivos específicos foram identificados para esse trabalho:

a) compreender a estrutura da rede;
b) compreender os protocolos utilizados para a transmissão de dados;
c) pesquisar e utilizar ferramentas para monitoramento do fluxo de dados da rede;
d) obter dados de tráfego a partir de amostragens na rede, utilizando algum software farejador;
e) compreender a utilização de métodos estatísticos que serão definidos durante o desenvolvimento;
f) analisar os dados obtidos e diagnosticar a rede conforme seu tráfego;
g) sugerir melhorias baseadas nos resultados.

1.3 JUSTIFICATIVA

As redes de computadores surgiram com objetivo de facilitar e/ou auxiliar na transferência e compartilhamento de informações. É necessário garantir aos usuários a disponibilidade de serviços adequados e com bom desempenho. A medida que as redes crescem, aumentam as dificuldades de seu gerenciamento, assim, utilizam-se ferramentas para sua monitoração e controle, que são essenciais para assegurar que sistemas de computadores funcionem corretamente e que interrupções não ocorram (KUROSE; ROSS, 2006).

Guillermo (2008) explica que os dados gerados por ferramentas de monitoração traçam um perfil de comportamento da rede, são úteis para a análise dos problemas potenciais e auxiliam para ter um diagnóstico mais preciso do uso dos recursos da rede. Para a segurança e o gerenciamento adequado das redes de computadores, é imprescindível que seu administrador tenha conhecimento de toda situação, identificando a presença de comportamentos indesejados, tais como: máquinas não autorizadas conectadas em rede *wifi*, consumo excessivo de banda, descoberta de vírus, entre outros.

Diante dessas informações, esta pesquisa visa analisar o que trafega pela rede de computadores da UNESC, por meio de uma metodologia de captura de informações na saída/entrada da rede. Para tal finalidade, pretende-se elaborar um padrão de coleta de dados utilizando-se ferramentas de monitoramento de redes (*sniffer*). Desse modo, é possível analisar os dados obtidos verificando seu propósito e se estão dentro das políticas de uso da universidade. Consequentemente, verifica-se a existência de decisões que possam ser tomadas para eliminar este tráfego indesejado ou possuir controle sobre ele. Com isso, obtém-se uma rede com melhor desempenho e vazão de informação.

1.4 ESTRUTURA DO TRABALHO

Demonstra-se suscintamente a seguir, o que cada capítulo deste trabalho irá abordar.

O primeiro capítulo fornece uma visão geral desta pesquisa, expondo os objetivos a serem alcançados.

O segundo capítulo conceitua como funciona uma rede de computadores e a comunicação entre os dispositivos, com a explicação dos principais protocolos de comunicação da arquitetura TCP/IP.

O terceiro capítulo apresenta as ferramentas de monitoramento, em especial os *sniffers*, detalhando o que são, seu funcionamento e suas principais características e funcionalidades.

O quarto capítulo aborda os métodos estatísticos empregados na execução e análise da pesquisa.

O quinto capítulo explana os trabalhos correlatos a esta pesquisa.

O sexto capítulo aborda de forma detalhada os processos para realização da análise do tráfego de rede da Universidade do Extremo Sul Catarinense, os métodos aplicados, os horários, períodos e tempo para cada amostragem. Ainda neste capítulo são demonstrados os resultados, as dificuldades e a conclusão da realização de todas as etapas.

2 REDES DE COMPUTADORES

Segundo Hayden (1999), o uso das redes vem se tornando um recurso indispensável em todos os locais onde exista um conjunto de computadores. Cada vez mais presentes no dia-a-dia das pessoas, estão em diversos locais: grandes e médias empresas, pequenos escritórios ou até mesmo em casa.

O objetivo de uma rede é realizar a comunicação entre dispositivos de características semelhantes por meio de regras, que assegurem a entrega da informação sem danificar os dados. Ou seja, vários dispositivos devem se reconhecer dentro de uma rede e serem capazes de transmitir as informações conhecendo seu destino (HAYDEN, 1999).

Uma rede de computadores pode ser definida como dispositivos (no mínimo dois) ligados de forma física (fio de cobre, fibras ópticas, ar, entre outros), com o intuito de compartilhar recursos físicos e lógicos (SOARES; LEMOS; COLCHER, 1995).

Com o crescimento significativo das redes de computadores, ela tornou-se indispensável para muitos ramos de negócios. O principal motivo dessa expansão se deve ao fato de sua utilização por empresas que compartilham informações, por usuários domésticos para diversão e entretenimento e até mesmo em instituições de ensino, onde alunos e professores conseguem informações do mundo todo, através das bibliotecas on-line (COMER, 2006).

Segundo Tanenbaum (2011) o objetivo das redes de computadores é deixar todos os recursos ao alcance das pessoas, independentemente de onde esteja localizado fisicamente o usuário e o meio a ser utilizado. Destaca-se um exemplo comum no dia a dia, que é o uso de uma impressora em um escritório com vários funcionários, que por meio da ligação em rede elimina-se a necessidade de ter uma impressora individual para cada usuário, consequentemente reduzindo o custo.

2.1 REDES LOCAIS E REDES DE LONGA DISTÂNCIA

As redes locais ou Local Area Network (LAN), são conhecidas por realizarem a conexão entre equipamentos num edifício de uma mesma organização

ou num campus universitário, com limitações de alguns quilômetros de distância. Normalmente, a tecnologia utilizada pelas LANs é por meio de cabos metálicos (par trançado) que em sua forma mais simples, opera com velocidade de transferência dos dados de 10Mbps a 100Mbps ou em LANs modernas podendo chegar até 10Gbps (TANEMBAUM, 2011).

Dentre as características das LANs, destaca-se que são de propriedade privada, possuem alta velocidade, baixo custo, confiabilidade, facilidade de acesso, padronização, entre outros (SOARES; LEMOS; COLCHER, 1995).

Com base nessas informações, observa-se que independente da literatura (tanto as atuais como as antigas) as descrições de uma LAN continuam as mesmas, o que mudou com o passar do tempo foram as tecnologias.

Quando as ligações de uma rede começam a tomar proporções metropolitanas, ou seja, quando uma empresa possui mais de um escritório na mesma cidade, e deseja que os computadores continuem interligados, então não se identifica mais como rede LAN e sim como Metropolitan Area Network (MAN), na qual mantem praticamente todas as características das redes locais, porém, torna-se possível a conexão dessas LANs dentro de algumas dezenas de quilômetros (SOARES; LEMOS; COLCHER, 1995).

Uma rede de longa distância Wide Area Network (WAN), é semelhante as LANs cabeadas, ela realiza a ligação de vários sistemas de computadores, porém, em distâncias maiores que as MANs (TANENBAUM, 2011).

Segundo Comer (2006) as WAN apesar de realizarem a comunicação entre longas distâncias (países ou continentes diferentes), sofrem com o quesito velocidade, pois não conseguem operar com altas taxas de conexão, podendo variar 1,5 Mbps até 2,4 Gbps.

Conforme relata Forouzan (2006) para o funcionamento de uma rede de computadores, seja ela LAN, MAN ou WAN, existe uma combinação de hardware (todo meio físico, infraestrutura e equipamentos), e software, que viabiliza todo o processo de comunicação. Existem vários processos a serem realizados ao se enviar um sinal de um computador para outro.

Comer (2015) elucida que para simplificar o projeto e a implementação de um protocolo**,** a comunicação é desmembrada em subproblemas que podem ser

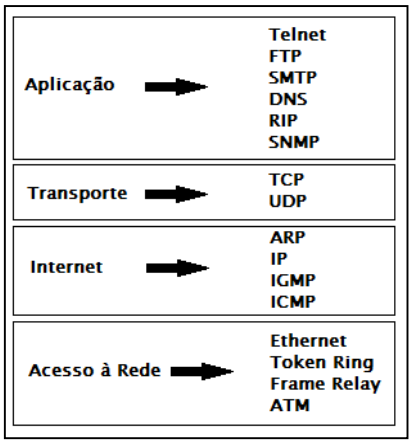
resolvidos de forma isolada. A ideia de camadas é crucial, pois permite dividir o problema em partes gerenciáveis, conforme demonstrado a seguir.

2.2 ARQUITETURA EM CAMADAS – TCP/IP

O TCP/IP ou pilha de protocolos, domina o processo de comunicação de dados e a conectividade de um computador com outro em uma rede. É conhecido desta forma devido aos seus dois principais protocolos, o Transmission Control Protocol (TCP) e o Internet Protocol (IP) (TANENBAUM, 2011).

Projetistas de rede fragmentaram o processo de transmissão de dados agrupando as funções que possuem características semelhantes e que dependem das outras, separando-as em camadas que recebem os agrupamentos de funções, distintas de outras camadas (FOROUZAN, 2006).

Figura 1 - Arquitetura TCP/IP



Fonte: Adaptado de Tanenbaum (2011).

Tanenbaum (2011) explica que a arquitetura TCP/IP é composta por quatro camadas que formam a pilha da estrutura do protocolo (Figura 1), da seguinte maneira:

a) camada de aplicação: onde ficam todos os programas do usuário e os protocolos de alto nível, como o de terminal virtual (TELNET), o de transferência de arquivos (FTP) e o de correio eletrônico (SMTP/POP3);
b) camada de transporte: responsável pela conectividade de dois hosts, mantendo a comunicação entre eles. Fazem parte desta camada os

protocolos Transmission Control Protocol (TCP) e User Datagram Protocol (UDP);

c) camada de Internet (camada de rede): responsável por fazer com que os hosts coloquem seus pacotes na rede e que cheguem até seu destino. Esta camada define um formato de pacote oficial e o Internet Protocol (IP), e mais um protocolo que o acompanha o Internet Control Message Protocol (ICMP);
d) camada de enlace: responsável por encapsular os pacotes da camada Internet no formato específico da rede associada, extrair os pacotes dos quadros vindos da rede (camada física) que não são confiáveis, e encaminha para camada Internet (camada de rede), corrigindo as possíveis falhas antes de encaminhar;
e) camada física: responsável por transmitir os bits como sinais eletrônicos através de meios físicos. Define as características mecânicas e elétricas da interface entre o dispositivo que transmite e os meios de transmissão.

Neste contexto, Farrel (2005) explica que os protocolos atendem a quatro propósitos. Codificam e transferem dados de um ponto para outro, e para tal, podem ter que controlar o modo de como os dados são distribuídos, designando os caminhos que precisam seguir. Para isso, pode ser necessária a troca de informações de estado de rede, e por fim, podem ser necessários para gerenciar os recursos de rede para controlar seu comportamento. Nesse sentido, aborda-se a seguir os protocolos de comunicação.

2.3 PROTOCOLOS DE COMUNICAÇÃO

O protocolo de comunicação tem a função de realizar a comunicação entre duas ou mais entidades (sistemas de gerenciamento de banco de dados, pacotes de transferência de arquivos, terminais, entre outros) e em sistemas diferentes (computadores, terminais e sensores remotos) (KUROSE; ROSS, 2010).

A seguir serão abordados os protocolos com maior incidência no cenário proposto, obtidos através de uma das ferramentas detalhadas posteriormente.

2.3.1 Protocolo IP

O Internet Protocol (IP) foi projetado com o intuito de realizar a interligação de redes, é o protocolo que mantém a Internet unida (TANENBAUM, 2011).

É o responsável pelo endereçamento e encaminhamento do pacote que será transmitido na rede até chegar ao seu destino. Diferente do TCP, ele não controla a conexão entre a origem e o destino do protocolo, apenas envia o pacote IP pela rede, fazendo seu roteamento e encaminhando ao seu destino, conforme seu endereço IP (SOUZA, 2007).

Figura 2 - Datagrama do protocolo IP

Tipo e versão	Tamanho do data-grama	Identifi-cação	Controle de fragmen-tação	Controle de tempo	Identif. protocolo superior	CRC	Endereço IP origem	Endereço IP destino	Opções IP	Dados
2 bytes	2 bytes	2 bytes	2 bytes	1 byte	1 byte	2 bytes	4 bytes	4 bytes	4 bytes	(variável)

Fonte: Adaptado de Souza (2007).

Conforme Tanenbaum (2011) descreve-se cada campo de controle e sua função da seguinte maneira:

a) tipo e versão: controla a versão do protocolo ao qual o datagrama pertence;
b) tamanho do datagrama: tamanho total do datagrama, possui este cabeçalho pois o tamanho não é constante;
c) identificação: identifica o datagrama, onde o *host* destino determina qual datagrama o fragmento recém chegado pertence;
d) controle de fragmentação: usado se houver fragmentação do datagrama na rede, informa que ponto o fragmento do datagrama pertence;
e) controle de tempo: é utilizado para controlar o tempo de vida útil dos pacotes na rede;
f) identificação do protocolo superior: identifica a qual protocolo o datagrama está transportando os dados;
g) CRC: controle de erros;
h) endereço IP de origem: endereço IP do transmissor;

i) endereço IP de desitno: endereço IP do receptor;

j) opções: usado em testes, segurança e análise de erros na rede;

k) dados: dados das camadas superiores levados pelo datagrama IP.

As características que se destacam no protocolo IP, é que a entrega do pacote não é garantida, podendo ocorrer perdas durante a transmissão. Podem haver atrasos, chegarem desordenados e até mesmo seguirem rotas ou caminhos diferenciados (SOUZA, 2007).

Atualmente existem duas versões do IP, a quatro ou IPv4 (mais utilizada atualmente) e a seis ou IPv6 (desenvolvida para substituir o IPv4).

2.3.2 Protocolo IP versão 4 (IPv4)

O IPv4 trata-se de um protocolo com 32 *bits* subdivididos em quatro grupos de 8 *bits* cada. O endereço IP por convenção é escrito com quatro inteiros decimais que são os grupos de oito *bits*, separados por pontos. Cada inteiro forma um valor decimal contido em oito *bits* de endereço (PETERSON; DAVIE, 2004).

Segundo Gallo e Hancock (2003), o IPv4 está dividido em cinco classes que são determinadas pelos quatro primeiros *bits* de seus 32 *bits*:

a) endereços Classe A: se o primeiro *bit* for igual a 0, o endereço é da Classe A tendo seu início em 0 a 127, ou de forma binária seu início é 00000000 a 01111111 no primeiro octeto;

b) endereços Classe B: se os dois primeiros *bits* forem iguais a 10, o endereço é da Classe B tendo seu início em 128 a 191, ou de forma binária seu início é 10000000 a 10111111 no primeiro octeto;

c) endereços Classe C: se os três primeiros *bits* forem iguais a 110, o endereço é da Classe C tendo seu início em 192 a 223, ou de forma binária seu início é 11000000 a 11011111 no primeiro octeto;

d) endereços Classe D: se os quatro primeiros *bits* forem iguais a 1110, o endereço é da Classe D tendo seu início em 224 a 239, ou de forma binária seu início é 11100000 a 11101111, porém, estes endereços classe D são utilizados apenas para *multicast*;

e) endereços Classe E: se os quatro primeiros *bits* forem iguais a 1111, o endereço é da Classe E tendo início 240 a 255. Todos são reservados para utilização futura.

Tanenbaum (2011) afirma que o endereço IP não é um *host* e sim uma interface de rede. Desse modo, se um *host* estiver em duas redes diferentes ele terá que possuir dois endereços IP e na maioria das vezes o *host* está ligado em apenas uma rede, com apenas um endereço IP.

2.3.3 Protocolo IP versão 6 (IPv6)

Com o crescimento da Internet, os números de endereços do IPv4 começaram a ficar escassos. Houve aumento também do número de novas redes, e como consequência, as tabelas de roteamento estão ficando lotadas e incapazes de lidar com a demanda de manter informações de todas as redes (GALLO; HANCOCK, 2003).

O IPv6 foi elaborado para substituir o IPv4, pois ele utiliza endereços com 128 *bits*, o que significa que temos 2^{128} endereços IPv6 contra 2^{32} endereços IPv4. Porém, apesar de ter grande semelhança com o IPV4, ele tem difícil implementação por ser diferente da camada de rede e não interligar com o IPv4 (TANENBAUM, 2011).

Tanenbaum (2011) destaca que mesmo o IPv6 tendo incompatibilidade com o IPv4, ele tem afinidade com protocolos auxiliares da Internet, incluindo TCP, UDP, ICMP, IGMP e DNS, sendo necessário apenas realizar alterações para que possam manipular endereços mais longos.

Além do tamanho do endereço do IPv6, outra melhoria a se destacar é referente ao seu cabeçalho, por ser mais simples, possuindo sete campos, ao invés do IPv4 que são treze. Com este aperfeiçoamento, os roteadores processam os pacotes com mais agilidade, melhorando o *throughput* e o atraso (TANEMBAUM, 2011).

Kurose (2010) explica que o IPv6 tem várias vantagens sobre o IPv4. Os dispositivos novos que já tem suporte ao IPv6 conseguem enviar, rotear e receber pacotes IPv4, contudo, é válido ressaltar que os dispositivos com suporte ao IPv4 não podem manusear pacotes IPv6.

2.3.4 Protocolo Internet Control Message Protocol (ICMP)

O protocolo ICMP é responsável por relatar ao transmissor quando acontece algo inesperado durante o processamento de algum pacote no roteador. Também é utilizado para efetuar a verificação da Internet, onde as mensagens são armazenadas e transportadas dentro de um pacote IP (TANENBAUM, 2011).

Os principais tipos de mensagens de erros do ICMP segundo Tanenbaum (2011) são:

a) destino não alcançável: é utilizado quando o roteador não consegue localizar o destino;
b) tempo excedido: é utilizado quando o Time To Live (TTL) do pacote chegou a zero, e ele foi descartado;
c) problema de parâmetro: é utilizado quando um valor inválido foi detectado no campo de cabeçalho;
d) redirecionamento: é utilizado quando é detectado que um pacote parece estar roteado incorretamente. É um aviso para o transmissor alterar sua rota por uma melhor;
e) pedido de informações e respostas de informações: é utilizado pelo *host* para verificar se o destino pode ser alcançado e se está ativo.

Além dessas mensagens de erros em destaque, existe uma lista mantida online em http://www.iana.org/assignments/icmp-parameters.

2.3.5 Protocolo Address Resolution Protocol (ARP)

O protocolo ARP tem uma grande importância em relação aos protocolos da camada de Internet. Ele é responsável por fazer o reconhecimento do endereço físico ou Media Access Control (MAC) de uma placa de interface de rede ou Network Interface Cards (NICs) em uma rede local, correspondente a um endereço IP (Souza 2007).

O protocolo ARP permite a comunicação entre duas ou mais entidades por meio da rede mesmo quando somente o endereço IP é conhecido pelo destinatário.

Para obter o endereço físico (MAC) da entidade destino, o protocolo ARP envia um broadcast com o IP do destinatário solicitando-o. Todas as entidades recebem a mensagem e as analisam, mas somente a que conter o IP correspondente responderá com o seu endereço MAC (COMER, 2006).

2.3.6 Protocolo Reverse Address Resolution Protocol (RARP)

O protocolo RARP executa exatamente o oposto do protocolo ARP visto anteriormente. A partir do endereço físico (MAC) de uma placa de interface de rede (Network Interface Card - NIC) conhecido do destinatário, determina o endereço da rede IP correspondente (SOUZA, 2007).

O protocolo RARP não tem tanto destaque na camada Internet, mas já foi muito utilizado para fazer boot de sistemas que não possuíam armazenamento estável. Seu funcionamento é simples, pois o sistema transmite um pedido RARP para receber o endereço IP do emissor (COMER, 2015).

2.3.7 Protocolo TCP

O Protocolo de Controle de Transmissão, do inglês Transmission Control Protocol (TCP), é responsável por garantir que os blocos de dados que estão sendo transmitidos cheguem sem falhas à sua aplicação destino (STALLINGS, 2005).

O TCP recupera os erros ocorridos na transmissão e solicita a retransmissão dos pacotes que foram perdidos ou eliminados pela rede, garantindo assim a integridade dos dados que são enviados e recebidos. Esse reenvio ocorre a partir da falha encontrada ou do pacote ausente, pois o TCP controla a recepção dos pacotes no seu destino garantindo sua recepção sem falhas (SOUZA, 2007).

Os pacotes chegam ao destino de forma aleatória, podendo chegar em ordem trocada das que foram enviadas da entidade emissora, e o TCP fica responsável por reorganizar os pacotes em mensagens na forma correta (TANENBAUM, 2011).

O TCP prepara o transporte de um transmissor para um receptor (fim a fim), criando uma conexão lógica e confiável entre eles, realizando o controle dos pacotes

enviados e recebidos, detectando eventuais perdas durante a transmissão, e informando sobre o recebimento dos mesmos (SOUZA, 2007).

Figura 3 - Formato de pacote do protocolo TCP com seus campos de controle e dados que transporta para aplicação

Port de origem	*Port* de destino	Número de seqüência	Confirmação de recebimento	Tamanho do header e bits de código	Tamanho da janela de recebimento	CRC controle de erros	Indicador de urgência	Opções	Dados
2 bytes	2 bytes	4 bytes	4 bytes	2 bytes	2 bytes	2 bytes	2 bytes	4 bytes	Variável

Fonte: Adaptado de Souza (2007).

De acordo com Souza (2007) descreve-se cada campo de controle e sua função do seguinte modo:

a) porta de origem: identifica o número da porta da aplicação do transmissor que fez a chamada;
b) porta de destino: identifica o número da porta da aplicação do receptor que é chamada;
c) número de sequência: número do segmento transmitido utilizado para garantir a sequência correta de chegada dos segmentos;
d) confirmação de recebimento: confirmação do segmento recebido por meio do envio do número do próximo *byte* esperado;
e) tamanho do *header:* indica o tamanho dos campos de controle (quatro *bits*), seis *bits* reservados e mais seis *bits* de códigos de controle de estabelecimento e encerramento de sessões de comunicação;
f) tamanho da janela: indica o número de pacotes que o receptor recebe antes de fazer a confirmação de recebimento;
g) CRC: controle de erros (*checksum*) calculado do cabeçalho e dos campos de dados;
h) indicador de urgência: indica se o pacote deve ter prioridade na transmissão;
i) opções: controles a serem definidos;
j) dados: dados vindos das camadas superiores e que serão transportados.

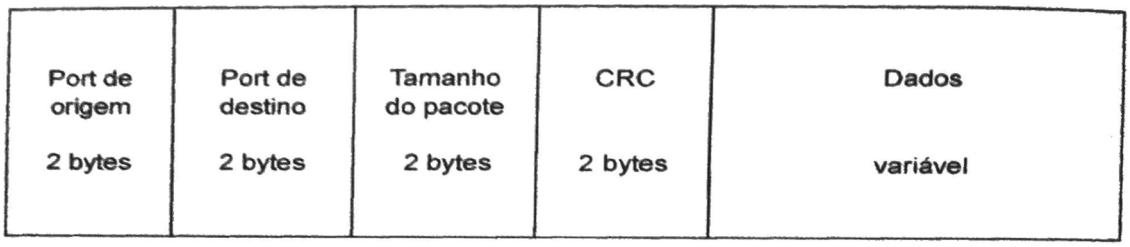
O TCP é o protocolo de transporte mais conhecido. Contudo, existem outros, como o User Datagram Protocol (UDP) que não realiza a conexão fim a fim, o que não garante a integridade dos dados que são emitidos, pois ele não executa a verificação da ausência de pacotes, a sequência de envio e se eles foram recebidos ou não (SOUZA, 2007).

2.3.8 Protocolo UDP

O User Datagram Protocol (UDP) segundo Gallo e Hancock (2003) não realiza a detecção ou correção de erros entre pontos terminais de transmissão, não retransmite dados, não trata erros, não controla o fluxo e não recebe confirmação de recebimento. Em outras palavras, toda a parte de credibilidade e segurança devem ser fornecidos pelos programas de aplicação que usam o UDP.

Por não realizar o tratamento dos controles citados anteriormente, o UDP se torna mais simples e rápido que o TCP. A diferença básica entre eles é que o TCP é um protocolo que inclui vários mecanismos para iniciar, manter e encerrar a comunicação, negociar tamanhos de pacotes, detectar e corrigir erros, evitar congestionamento do fluxo e permitir a retransmissão de pacotes corrompidos (SOUZA, 2007).

Figura 4 - Formato do pacote UDP



Fonte: Adaptado de Souza (2007).

De acordo com Gallo e Hancock (2003) descreve-se cada campo de controle e sua função do seguinte modo:

a) porta de origem: identifica a aplicação do transmissor que está chamando;
b) porta de destino: identifica a aplicação no receptor que está sendo chamado;

c) tamanho do pacote: tamanho do pacote ou datagrama do UDP, incluindo campos e controle e dados;
d) CRC: controle de erros, quando não utilizado é preenchido com zeros;
e) dados: dados originários das camadas superiores criados por um protocolo de aplicação como o SMTP ou POP3 para e-mail que serão transportadas.

Em momentos em que a rede estiver congestionada, as aplicações que utilizam o UDP tendem a ter um menor desempenho, visto que ele não realiza as verificações como o TCP (GALLO; HANCOCK, 2003).

2.3.9 Protocolo File Transfer Protocol (FTP) e Trivial File Transfer Protocol (TFTP)

O FTP é o principal protocolo de transferência de arquivos entre computadores e responsável por grande parte do tráfego em uma rede TCP/IP. Ele proporciona interação entre o cliente e servidor, com segurança, por meio de identificações e senhas (SOUZA, 2007).

Segundo Kurose e Ross (2010) o FTP serve para efetuar transferência de arquivos entre dois computadores e utiliza duas conexões TCPs em paralelo, uma para controle e outra de dados. A primeira serve para enviar informações contendo as identificações, senhas e comandos para obter e inserir arquivos. Já a conexão de dados tem o propósito de enviar os arquivos.

O protocolo TFTP tem a mesma finalidade do FTP, com algumas características distintas. Uma delas é sua programação ser menos complicada por não ter interações complexas entre cliente e servidor e também por dispensar a autenticação por usuário e senha. É uma aplicação da arquitetura TCP/IP que serve para transferir arquivos entre dois computadores em uma rede IP utilizado com o UDP. Porém, sua transferência de arquivos não possui controles de fluxo e nem de sequência de pacotes, tornando este tipo de transmissão de arquivos mais rápida e sujeita a falhas na transmissão (SOUZA, 2007).

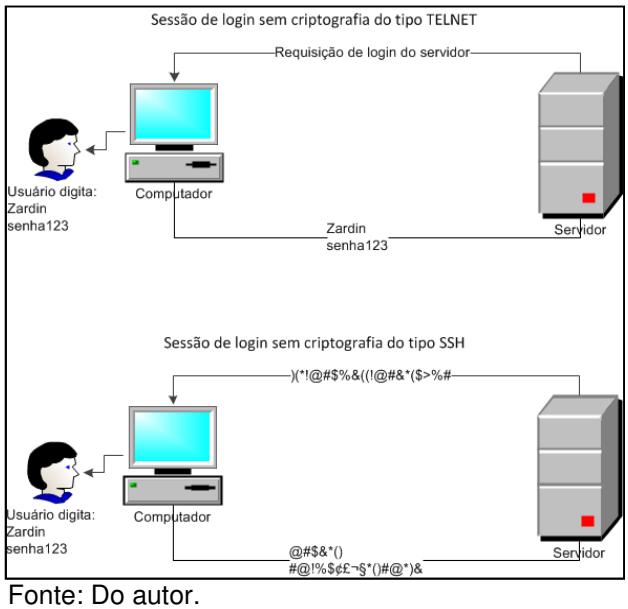
2.3.10 Protocolo Terminal Emulator (Telnet) e protocolo Secure SHell (SSH)

O TELNET é semelhante ao protocolo FTP e é responsável por parte do tráfego gerado em uma rede. É utilizado por administradores para o controle de dispositivos, ou para usuários que queiram um serviço de terminal virtual, ou seja, que necessitam de uma interação com uma máquina remota, como se estivessem na frente do computador (COMER, 2015).

Este protocolo fornece um mecanismo onde todos os caracteres digitados na máquina local, sejam repassados diretamente para a remota por meio do protocolo TCP, garantindo o transporte confiável dos dados. Ele é acessível por meio do programa de aplicação *telnet*, encontrado nos sistemas operacionais Linux, Windows e MAC OS (GALLO; HANCOCK, 2003).

O SSH é uma alternativa para o TELNET que permite ao usuário acesso virtual a algum servidor como se estivesse na frente dele, com a diferença de poder usar criptografia para manter as sessões confidenciais (COMER, 2015).

Figura 5 - Conexão do tipo TELNET e do tipo SSH



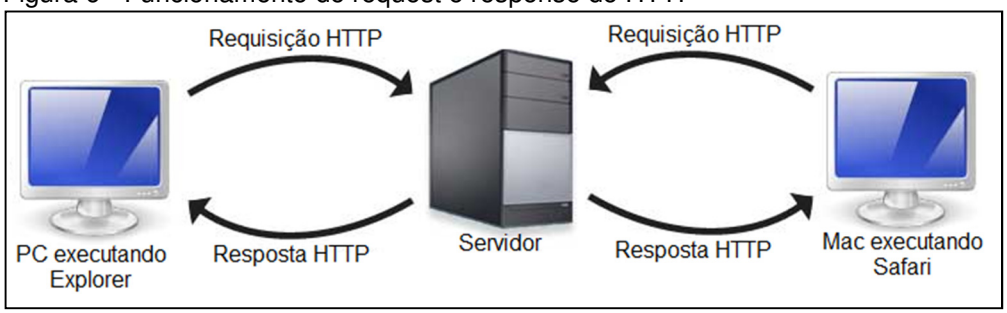
Fonte: Do autor.

2.3.11 Protocolo HiperText Transfer Protocol (HTTP)

O HTTP é executado sobre o protocolo TCP e por isso pertence a camada de aplicação. Permite a transferência de dados entre redes de computadores,

principalmente na World Wide Web (WWW). O uso mais comum deste protocolo é entre navegadores Web e um servidor Web, onde cada transação é tratada independente e para cada uma será criada uma nova conexão TCP entre o cliente e o servidor, posteriormente, quando esse processo finalizar, a conexão será encerrada (STALLINGS, 2005).

Figura 6 - Funcionamento do request e response do HTTP



Fonte: Adaptado de Kurose e Ross (2010).

Comer (2015) destaca as seguintes características do protocolo HTTP:

a) *request* / *response*: estabelece uma sessão de transporte que envia um pedido de HTTP para um servidor, que obtém resposta;
b) *stateless*: o HTTP considera cada requisição como uma transação independente, onde nenhuma requisição anterior ou sessão anterior influencia na transação atual;
c) transferência bidirecional: normalmente os navegadores solicitam uma página Web e o servidor transfere uma cópia para o navegador. O HTTP permite que aconteça o oposto, o navegador que transfere informações para o servidor;
d) capacidade de negociação: o HTTP permite que os navegadores e servidores especifiquem os recursos que podem utilizar durante a transferência e quais são aceitos;
e) suporte para o cache: refere-se ao armazenamento de uma cópia da página Web no navegador e se o usuário solicitar uma página, o navegador pode interromper o servidor para verificar se existe alguma mudança desde que ocorreu o armazenamento em cache;

f) suporte para os intermediários: o HTTP permite que uma máquina atue como um servidor proxy que armazena as páginas Web e forneça respostas as solicitações de um navegador através do seu cache.

2.3.12 Correio eletrônico (e-mail)

Este serviço é um dos responsáveis pelo grande tráfego de Internet numa rede. Com suas melhorias e evolução, foram transformadas em aplicações Web onde o usuário acessa seu e-mail e tem a opção de verificar suas mensagens, respondê-las, encaminhá-las ou selecionar uma dessas mensagens para tomar uma ação. Quando o usuário envia uma mensagem, o sistema utiliza protocolos de transferência de e-mail que se encarrega de transportá-la até seu destinatário (COMER, 2015).

2.3.12.1 Simple Mail Transfer Protocol (SMTP)

O protocolo SMTP é responsável pelo envio e recepção do e-mail, ou seja, ele realiza a comunicação entre computadores e servidores de correio eletrônico. As mensagens são entregues quando o computador que está enviando (cliente SMTP) estabelece uma conexão TCP com a porta 25 do computador de destino (servidor SMTP), e se o destinatário estiver inoperante ou indisponível, o cliente tenta entregar a mensagem posteriormente (KUROSE; ROSS, 2010).

O protocolo SMTP é o mecanismo padrão de correio eletrônico na Internet e é mais simples do que o protocolo que o antecede, o Mail Transfer Protocol (MTP). Como vários protocolos da camada de aplicação, o SMTP realiza a comunicação entre cliente e servidor utilizando um texto ASCII que torna fácil o seu desenvolvimento, testes e depuração (COMER, 2015).

Segundo Tanenbaum (2011) esta comunicação entre cliente e servidor é realizada com comandos abreviados, cita-se como exemplo, o HELO (abreviação de HELLO), ou com números contendo três dígitos.

Figura 7 - Transferência de mensagens entre Alice e Bob

```
                S: 220 ee.uwa.edu.au SMTP service ready
C: HELO cs.washington.edu
                S: 250 cs.washington.edu says hello to ee.uwa.edu.au
C: MAIL FROM: <alice@cs.washington.edu>
                S: 250 sender ok
C: RCPT TO: <bob@ee.uwa.edu.au>
                S: 250 recipient ok
C: DATA
                S: 354 Send mail; end with "." on a line by itself
C: From: alice@cs.washington.edu
C: To: bob@ee.uwa.edu.au
C: MIME-Version: 1.0
C: Message-Id: <0704760941.AA00747@ee.uwa.edu.au>
C: Content-Type: multipart/alternative; boundary=qwertyuiopasdfghjklzxcvbnm
C: Subject: A Föld egész számú alkalommal kerüli meg a Napot
C:
C: Ez a bevezető. A felhasználói ügynök figyelmen kívül hagyja. Minden jót.
C:
C: --qwertyuiopasdfghjklzxcvbnm
C: Content-Type: text/html
C:
C: <p> Boldog szülinapot<br>
C: Boldog szülinapot<br>
C: Boldog szülinapot, kedves <b> Bob </b><br>
C: Boldog szülinapot</p>
C:
C: --qwertyuiopasdfghjklzxcvbnm
C: Content-Type: message/external-body;
C:       access-type="anon-ftp";
C:       site="bicycle.cs.washington.edu";
C:       directory="pub";
C:       name="birthday.snd"
C:
C: content-type: audio/basic
C: content-transfer-encoding: base64
C: --qwertyuiopasdfghjklzxcvbnm
C: .
                S: 250 message accepted
C: QUIT
                S: 221 ee.uwa.edu.au closing connection
```

Fonte: Adaptado de Tanenbaum (2011).

Como demonstrado na Figura 7, as linhas enviadas pelo cliente (o que está enviando o e-mail) são marcados com “C:”, e as linhas enviadas pelo servidor são marcadas com “S:”. O transmissor começa aguardando o comando “220” do servidor, e em seguida envia o comando HELO, como informado anteriormente é a abreviação de HELLO. Se o servidor estiver sobrecarregado, o cliente aguarda diminuir a quantidade de processos e depois inicia os próximos comandos para enviar a mensagem (TANENBAUM, 2011).

2.3.12.2 Post Office Protocol, version 3 (POP3)

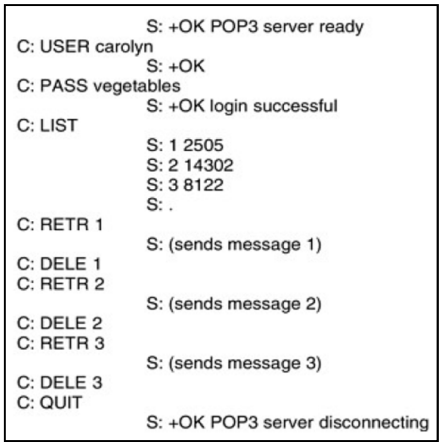
O POP3 é um protocolo popular para transferência de e-mails de um servidor de correio eletrônico para um computador local ou dispositivos móveis. Com isso, se enquadra como um dos vilões para o tráfego da rede.

O POP3 é um protocolo simples e funciona da seguinte maneira: o usuário utiliza uma aplicação cliente (POP3), que cria uma conexão TCP com o servidor de

correio eletrônico (POP3), informa um usuário e senha para estabelecer a conexão que quando aceita, são enviados os comandos para obter-se uma cópia dos e-mails da caixa de entrada para um computador local ou dispositivo móvel, sendo a original existente no servidor removida (COMER, 2015).

Na Figura 8 é demonstrado o funcionamento do processo entre cliente e servidor, depois da autenticação com o usuário e senha.

Figura 8 - Usando POP3 para baixar três mensagens



Fonte: Adaptado de Tanenbaum (2003).

O primeiro passo realizado pelo cliente após sua autorização é o envio do comando LIST, que faz o servidor retornar o conteúdo da caixa de correio (uma mensagem por linha e com seu tamanho). Ao final da lista é encerrado com ponto final “S: .”. Em seguida as mensagens são recuperadas com o comando RETR e logo deletadas do servidor com o comando DELE. Após todas as mensagens estiverem recuperadas e marcadas para serem excluídas do servidor, o cliente envia o comando QUIT, encerrando a transação e entrando em modo de atualização. Logo que o servidor terminar a exclusão das mensagens ele retorna uma resposta “S: +OK POP3 *server disconnecting*”, e interrompe a conexão TCP (TANENBAUM, 2003).

Uma desvantagem de realizar o download de e-mails para um computador local ou dispositivo móvel é pelo fato de não existir uma cópia na caixa de correio. Caso ocorra algum problema no computador local, pode perder todos os e-mails permanentemente. Outro fato a considerar, é que pode ser necessário acessar algum e-mail que já foi baixado para o computador local e não será possível. O protocolo POP3 permite deixar uma cópia das mensagens no servidor ao realizar o download,

contudo, a maioria das aplicações de correio eletrônico simplesmente baixa tudo e esvazia a caixa de correio. Nesse sentido, pode-se observar que este protocolo é bom para o servidor, e nem tanto para o usuário (TANENBAUM, 2011).

2.3.12.3 Internet Message Access Protocol (IMAP)

O protocolo IMAP surgiu com o intuito de suprir a necessidade de usuários que acessavam o e-mail de vários computadores e que, por organização ou necessidade, utilizavam pastas para melhor visualização e facilidade para localização dos e-mails. O POP3 apenas baixava os e-mails do servidor, e possuía a opção de manter uma cópia no servidor, foi então que para resolver esse e outros problemas surgiu o IMAP, que também é utilizado para acessar o correio eletrônico, porém, com mais recursos e complexidade. O IMAP versão 4 (IMAP4) seria uma melhoria ou alternativa para o protocolo mais antigo de entrega final, o POP3, que foi abordado anteriormente (KUROSE; ROSS, 2010).

Uma outra vantagem do protocolo IMAP sobre o POP3 é que o usuário pode procurar uma sequência de caracteres, ou obter apenas o cabeçalho de uma mensagem, tornando-se muito eficiente quando a largura de banda entre o cliente e o servidor for pequena ou a conexão for de baixa velocidade. Quando o usuário se deparar com esta situação, ele decide se baixa ou não todas as mensagens do servidor, evitando baixar mensagens longas ou que contenham arquivos grandes que não sejam de seu interesse (COMER, 2015).

Uma rede é de crucial importância para qualquer negócio, portanto, monitorar e otimizar seu desempenho é essencial. Destaca-se nos próximos capítulos os *sniffers*, softwares que auxiliam na identificação de problemas e serão a ferramenta chave na elaboração desta pesquisa.

3 FERRAMENTAS DE MONITORAMENTO

As redes de computadores são compostas por vários equipamentos que se comunicam e compartilham dados e recursos. Existem casos, que para uma tarefa ser bem executada por uma aplicação, a comunicação por meio da rede deve estar dentro de certos limites de desempenho (STALLINGS, 2005).

Além de aplicações e protocolos que atuam no nível da camada de rede e utilizam estes serviços de comunicação e compartilhamento de dados e recursos, um sistema/ferramenta é necessário para realizar o controle sobre estes recursos, identificar problemas e computadores que violam as políticas de uso de uma instituição (COMER, 2015).

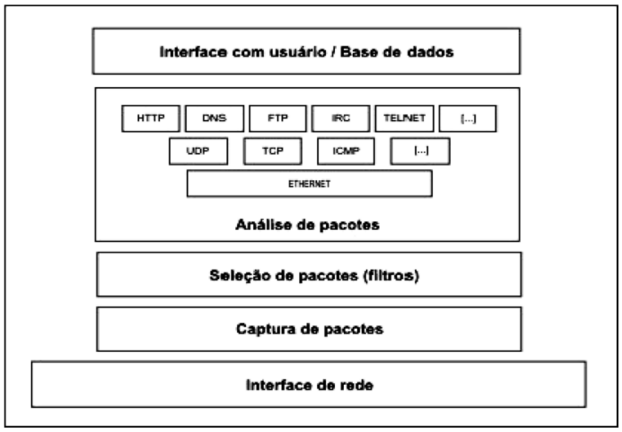
3.1 *SNIFFER*

Uma rede padrão Ethernet envia um pacote para todos os computadores que estão no mesmo segmento de rede e o cabeçalho deste pacote possui o endereço do computador destino onde somente ele receberia este pacote. Porém, existe a possibilidade de um computador capturar os pacotes mesmo que não sejam destinados a ele. Este computador que é colocado como escuta de todos os pacotes, diz-se que ele está em modo promíscuo (OLIVEIRA, 2000).

Os *sniffers* são dispositivos que realizam a captura de pacotes em uma rede, sendo sempre um conjunto de hardware e software, executando por meio de um programa ou aplicação *sniffer* e uma interface de rede em modo promíscuo. Independente do destino do pacote, a interface de rede do computador fará a leitura como se fosse destinada a ela. Essa captura de pacotes é nomeada de *sniffing*.

Segundo Furmankiewicz (2000) o termo *sniffer* é diferente de *Sniffer*, pois este com a inicial maiúscula foi um produto lançado em 1988 pela antiga Network General Corporation. Isto aconteceu devido a Network General Corporation ter dominado o mercado por muito tempo, tornando este termo popular e a partir daí os analisadores de protocolos que foram surgindo, ficaram conhecidos como *sniffer*.

Figura 9 - Arquitetura de um *sniffer*



Fonte: Jesus (2008).

A utilização do *sniffer* pode se tornar muito perigosa, pois por meio dele pode-se capturar senhas, informações confidenciais de proprietários, entre outros (FURMANKIEWICZ, 2000).

Com base nessas informações, a existência de um *sniffer* não autorizado em uma rede, indica que de fato todo sistema está comprometido e muitos problemas podem ocorrer. Devido a isso, o *sniffer* é associado a ferramentas de roubo de informações, ou má intencionada.

Por outro lado, o *sniffer* é um grande aliado dos administradores de rede, auxiliando na detecção de problemas, fornecendo informações de endereço de origem e destino, a formação dos pacotes e informações de alguns protocolos específicos que auxiliam na resolução de problemas (CASAGRANDE, 2003).

3.2 PRINCIPAIS *SNIFFERS* E SUAS FUNCIONALIDADES

Atualmente existem várias ferramentas do tipo *sniffer* com finalidades específicas, mas com as mesmas características. As versões pagas são muito eficientes e com muitos recursos. No entanto, existe ferramentas gratuitas e de código aberto que suprem as principais necessidades, tendo apenas como desvantagem o fato de não possuírem suporte.

Existem vários *sniffers*, alguns mais simples com recursos reduzidos, outros mais avançados sendo capazes de importar relatórios e arquivos de

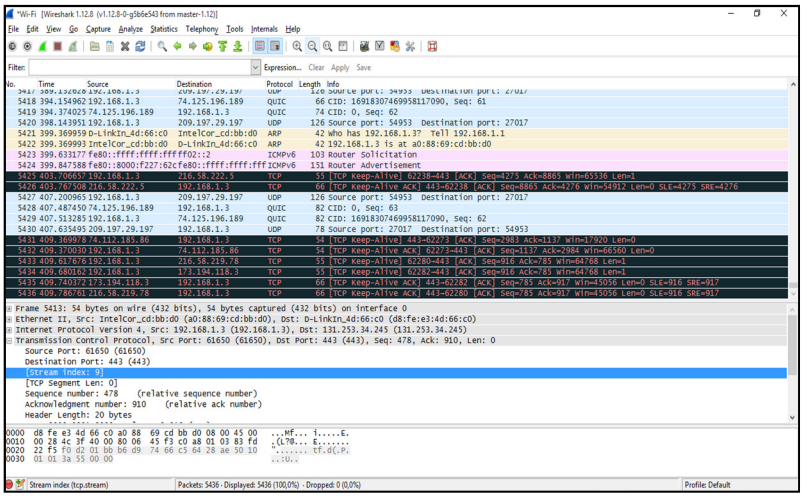
ferramentas de análise de rede. Contudo, um dos mais utilizados e conhecidos é o Wireshark (antigo Ethereal) (CASAGRANDE, 2003).

A seguir serão listados alguns *sniffers* e suas funcionalidades.

3.2.1 Wireshark

O Wireshark analisa os pacotes de rede, característica padrão dos *sniffers*. Ele executa a captura dos pacotes de rede e demonstra os dados destes pacotes de forma detalhada. É um *sniffer* de código aberto e está disponível para download em https://www.wireshark.org/download.html possui versão para os sistemas operacionais Windows, Linux e MAC.

Figura 10 - Interface gráfica do Wireshark



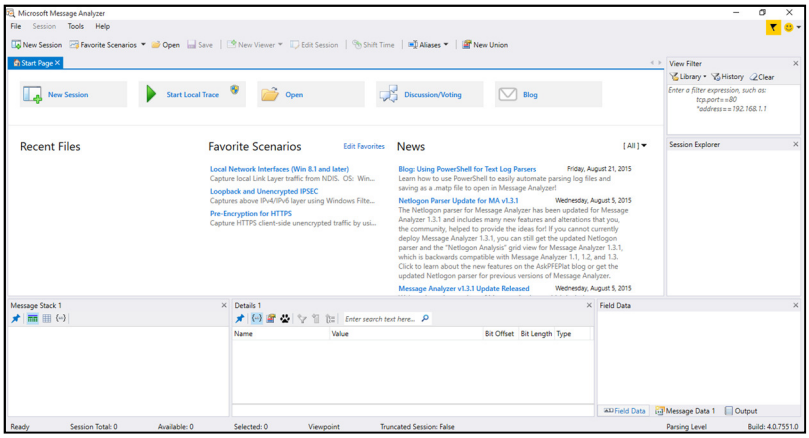
Fonte: Wireshark Foundation.

3.2.2 Microsoft Message Analyzer

Microsoft Message Analyzer trata-se do antigo Microsoft Network Monitor, realiza a análise, captura do tráfego de protocolos e outras mensagens do sistema.

Está disponível para download em http://www.microsoft.com/en-us/download/details.aspx?id=44226 e executa somente no sistema operacional Windows.

Figura 11 - Interface gráfica do Microsoft Message Analyzer

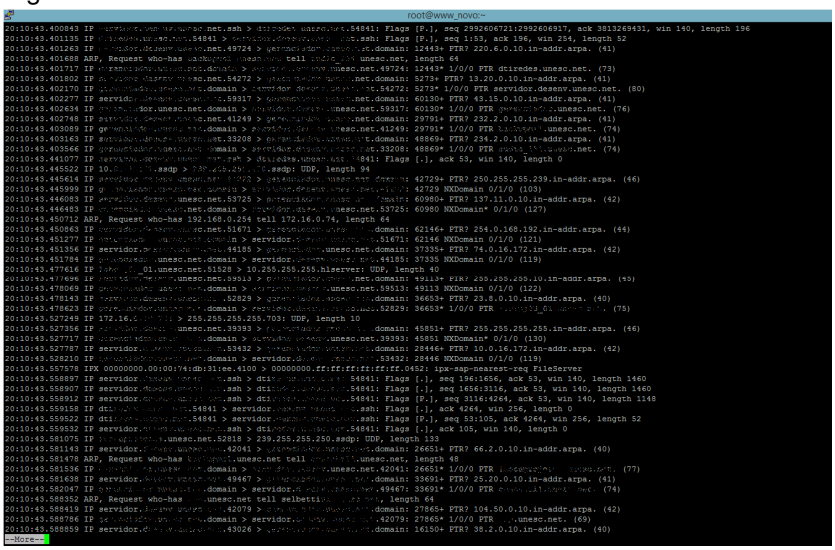


Fonte: Microsoft.

3.2.3 Tcpdump

O Tcpdump mostra o conteúdo dos pacotes, salva para análise posterior e também permite a leitura de arquivos salvos. Este é o mais antigo dentre os analisadores de pacotes de rede e está disponível para download em http://www.tcpdump.org/ em versões para Linux e MAC.

Figura 12 – TCPDUMP

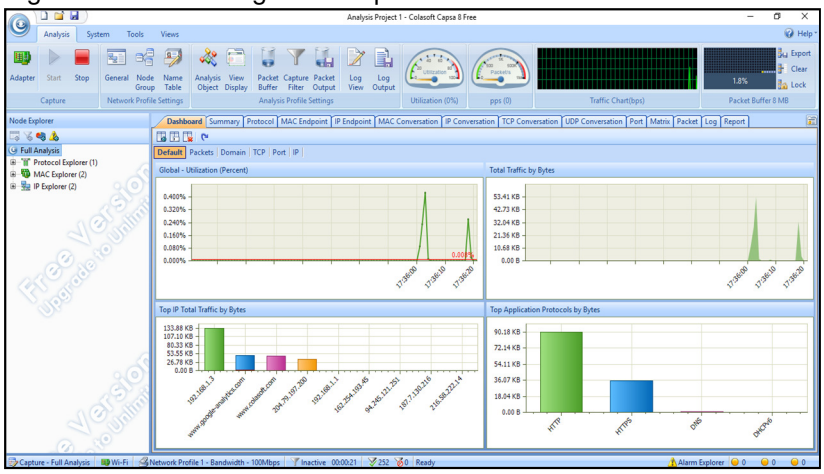


Fonte: LBNL's Network Research Group.

3.2.4 Capsa Packet *Sniffer*

Capsa Packet *Sniffer* é um analisador de rede voltado para a monitoramento, detecção de problemas, entre outros. Está disponível para download em http://www.colasoft.com/capsa-free/.

Figura 13 - Interface gráfica Capsa Packet *Sniffer*

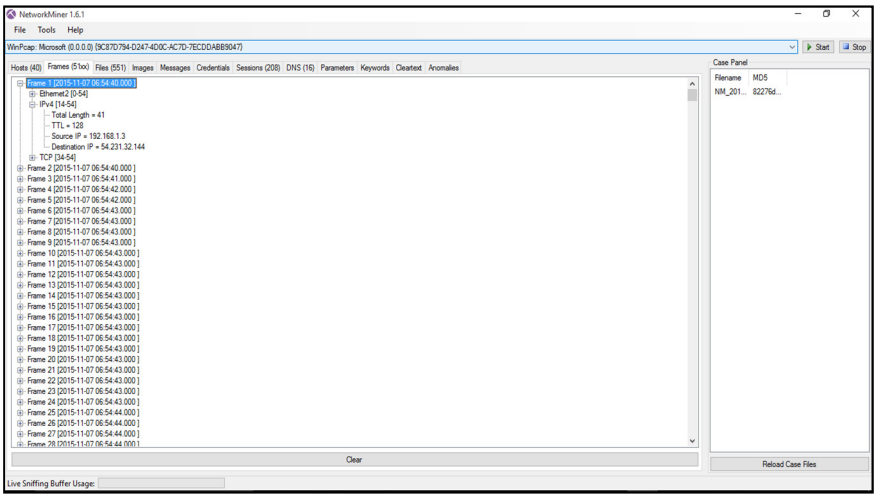


Fonte: Colasoft.

3.2.5 NetworkMiner

O NetworkMiner é uma ferramenta de análise forense de rede disponível para Windows, Linux, Mac e FreeBSD. Realiza a captura de pacotes na rede, com a finalidade de detectar sistemas operacionais, sessões, nome de computadores, portas abertas, dentre outras funcionalidades. O mais importante é não gerar nenhum tipo de tráfego na rede. Está disponível para download em http://www.netresec.com/?page=NetworkMiner.

Figura 14 - Interface gráfica do NetworkMiner



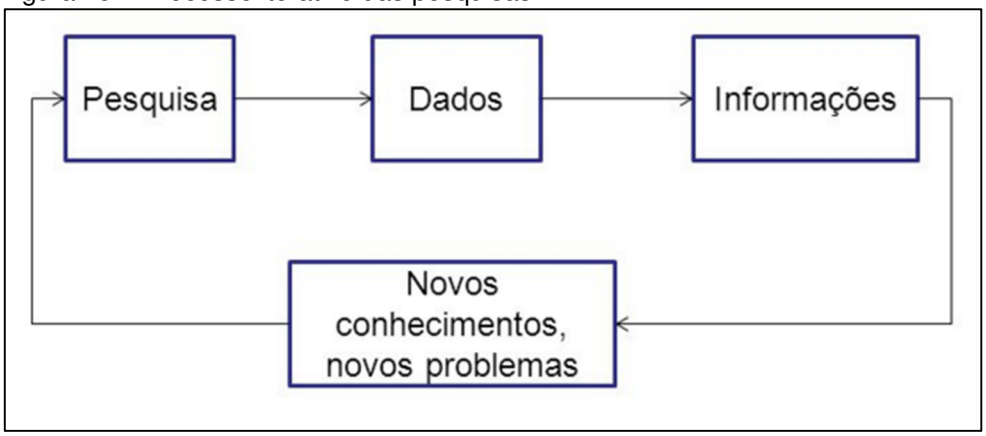
Fonte: Netresec.

Os softwares que foram apresentados anteriormente realizam a captura de pacotes que circulam em uma rede de computadores e com isso, geram um volume de dados muito grande devido a frequente troca de informações que acontece entre os dispositivos interligados via rede. Desse modo, o próximo capítulo aborda métodos estatísticos para auxiliar na coleta, análise, interpretação e apresentação dos dados gerados.

4 MÉTODOS ESTATÍSTICOS

A estatística é uma área da matemática que por meio dos números gera informações úteis para tomada de decisões, resolução de problemas ou produção de conhecimento. Contudo, a partir do momento que adquire-se novos conhecimentos, naturalmente surgem novos problemas e como consequência gera-se um processo iterativo (LEVINE et al, 2012).

Figura 15 – Processo iterativo das pesquisas



Fonte: Adaptado de Barbetta, Reis e Bornia (2010).

No campo da informática as informações são tratadas por meios eletrônicos e computacionais. Na estatística, obtém-se informações relevantes de massa de dados, normalmente com o auxílio de computadores (BARBETTA; REIS; BORNIA, 2010).

Aborda-se a seguir, os métodos que foram escolhidos por suas características, tendo como base os dados que foram obtidos durante a execução deste projeto, a partir das seguintes fases do método estatístico: coleta, crítica, apuração e apresentação dos dados. Finalizando-se o processo com a análise dos resultados obtidos.

4.1 TESTE U DE MANN-WHITNEY

Segundo Callegari (2004), é um teste não paramétrico para comparar a posição central entre duas populações, embasada em amostras independentes, extraídas de forma aleatória das populações, substituindo o teste *t* de Student.

Essa pesquisa visa demonstrar a comparação das médias de tráfego de rede com comportamento padrão, porém, em redes distintas, sendo uma administrativa e a outra de laboratório.

Para a realização do teste U deve-se primeiro realizar a ordenação das amostras, colocando-as em posto numerando-as do menor para o maior. Em seguida, determina-se a soma dos postos para cada amostra, que serão representadas por $R_1 \, e \, R_2$, na qual, $N_1 \, e \, N_2$ referem-se ao tamanho das amostras (SPIEGEL; STEPHENS, 2009).

Para o cálculo da estatística U aplica-se as fórmulas (1) e (2):

$$U = N_1 N_2 + \frac{N_1(N_1 + 1) - R_1}{2}$$

$$U = N_1 N_2 + \frac{N_2(N_2 + 1) - R_2}{2}$$

Com número elevado de amostras, o cálculo para aproximação normal z, é dada pela fórmula 3:

$$Z = \frac{U - \mu_U}{\sigma_U}$$

A fórmula demonstra que μ_U e σ_U representam o desvio padrão de U se a hipótese nula for verdadeira, dadas pelas fórmulas 4 e 5:

$$\mu_U = \frac{N_1 N_2}{2}$$

$$\sigma_U = \sqrt{\frac{N_1 N_2 (N_1 + N_2 + 1)}{12}}$$

4.2 TESTE H DE KRUSKAL-WALLIS

Trata-se de uma extensão do teste U de Mann-Whitney, é não paramétrico e substitui o teste ANOVA. Utilizado para a realização de comparações entre três ou mais populações (LEVINE et al, 2012).

Segundo definição de Spiegel e Stephens (2009), supondo-se a existência de k amostras com tamanhos $N_1, N_2 \ldots N_k$, ordena-se os dados em postos e as somas deles de k amostras são $R_1, R_2, \ldots R_K$, respectivamente. Aplicando-se a fórmula 6 $H =$

$\frac{12}{N(N+1)}\sum_{j=1}^{k}\frac{R_j^2}{N_j} - 3(N+1)$ observa-se que a distribuição amostral de H é próxima de uma qui-quadrado com k-1 graus de liberdade, desde que os tamanhos das amostras sejam iguais a 5.

4.3 METODO DE DUNN

É um teste não paramétrico utilizado para comparação de várias populações em pares. Quando ocorre uma rejeição da hipótese nula do teste de Kruskal-Wallis, é possível testar os pares e determinar se eles foram os responsáveis por essa rejeição no teste (ARANGO, 2005).

Para comparação dos grupos A e B no teste H com h grupos utiliza-se a fórmula 7.

$$z = \frac{\frac{U_A}{n_A} - \frac{U_B}{n_B}}{\sqrt{\frac{N.(N+1)}{12}.\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}}$$

A letra z representa a distribuição aproximadamente normal. Para demonstração do número de observações nos grupos A e B tem-se $n_A \, e \, n_B$.

4.4 TESTE KOLMOGOROV-SMIRNOV

É um teste não paramétrico para comparar a existência de diferença entre duas distribuições de probabilidades subjacentes. Para tal finalidade, existe a possibilidade de aplicação do teste qui-quadrado de aderência, porém, o teste de Kolmogorov-Smirnov demonstra ter melhores resultados em certas situações. Ele também é utilizado para verificar se a distribuição de variáveis quantitativas se aproxima de uma distribuição normal (BARBETTA, 2010).

A função de distribuição acumulada F_n para n observações y_i é definido pela fórmula 8:

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n}\begin{cases}1 & se\ y_i \leq x, \\ 0 & caso\ contrario.\end{cases}$$

As estatísticas de teste Kolmogorov-Smirnov é definido pelas fórmulas 9 e 10:

$$D_n^+ = \max(F_n(x) - F(x))$$

$$D_n^- = \max\big(F(x) - F_n(x)\big)$$

na qual, o F(x) representa a distribuição em hipótese e o D representa a diferença absoluta máxima.

5 TRABALHOS CORRELATOS

A análise de tráfego de rede auxilia os administradores de rede a identificarem problemas e saberem com mais detalhes o que trafega sobre ela. Existem vários estudos na comunidade científica abordando esse assunto. Ressalta-se a seguir os mais relevantes referentes a este trabalho.

5.1 DIAGNÓSTICO DO TRÁFEGO DE REDE DE LABORATÓRIOS DE INFORMÁTICA. ESTUDO DE CASO: UNIVERSIDADE DO EXTREMO SUL CATARINENSE

Trata-se de um Trabalho de Conclusão de Curso em Ciência da Computação desenvolvido na Universidade do Extremo Sul Catarinense – UNESC.

Trombim (2006) realizou uma análise de tráfego nos laboratórios da universidade, por meio do software Ethereal para Linux, que foi instalado no servidor proxy dos laboratórios e aplicada a técnica de amostragem estratificada para obter as amostras do tráfego. Notou-se com a realização do mesmo, a quantidade de pacotes que trafegava pela rede, observou-se o comportamento dos principais protocolos, aplicações e serviços encontrados, o tráfego e o horário de maior pico e o consumo dos recursos de rede.

Este trabalho apresenta semelhança com o que está sendo proposto nessa pesquisa, contudo, aponta-se como diferença mais significativa, o volume de dados e o número de equipamentos que serão analisados e também o maior número de resultados obtidos, pois a pesquisa será realizada em uma rede de laboratório e também na administrativa.

5.2 ANÁLISE DA REDE SOB O PONTO DE VISTA DO CONTROLE DE INFORMAÇÕES E TRÁFEGO ESTUDO DE CASO: TSA QUÍMICA DO BRASIL

Jesus (2008) realizou um Trabalho de Conclusão de Curso em Ciência da Computação na Universidade do Extremo Sul Catarinense – UNESC, com a análise no tráfego da rede da TSA Química do Brasil, na qual os administradores obtiveram

dados do que estava trafegando em sua rede, se existiam informações desnecessárias e as melhorias que poderiam ser realizadas na rede. Para análise foi utilizada a ferramenta Wireshark e também a técnica de amostragem estratificada proporcional.

5.3 DIAGNÓSTICO DO TRÁFEGO DE REDE WEB E ANÁLISE DA BASE DE DADOS GERADA PELO MICROSOFT INTERNET SECURITY AND ACCELERATION 2006: ESTUDO DE CASO NA SATC

Este Trabalho de Conclusão de Curso em Ciência da Computação foi desenvolvido na Universidade do Extremo Sul Catarinense – UNESC por Lucca (2009). Foi realizada uma análise do tráfego de rede na Escola técnica da SATC e por meio de relatórios gerados pela ferramenta Microsoft ISA Server 2006 e com o auxílio de métodos estatísticos, comprovou a lentidão da Internet em períodos de pico dentro da instituição. Com base nesses dados, foram apresentados diagnósticos para auxiliar na melhoria.

5.4 GERENCIAMENTO E MONITORAÇÃO DE REDES DE COMPUTADORES UTILIZANDO-SE ZABBIX

Monografia desenvolvida na Universidade Federal de Lavras – UFLA.

Bonomo (2006) realizou o estudo e a implantação de uma ferramenta que auxilia o administrador de rede a manter ela em funcionamento, minimizando erros por falta de controle. A ferramenta utilizada foi o Zabbix que é uma poderosa ferramenta no uso de gerenciamento e monitoramento de rede. Com esta pesquisa foi possível observar o poder de gerenciamento da ferramenta estudada e a importância dessas ferramentas para os administradores de rede.

6 ANÁLISE NO TRÁFEGO DA REDE ADMINISTRATIVA E DE LABORATÓRIO

Neste capítulo aborda-se de forma mais detalhada o desenvolvimento desta pesquisa, a definição dos pontos e horários de coleta por meio dos métodos estatísticos, bem como a utilização da ferramenta de captura de pacotes para obtenção dos dados e os resultados com a análise da amostragem.

Apresenta-se também os resultados obtidos a partir dos dados analisados, sobre o comportamento do tráfego nas redes verificadas.

6.1 METODOLOGIA DE ANÁLISE

Segundo Crespo (2009) quando necessita-se de uma conclusão sobre um todo (população), é normal observar-se esse todo na forma de amostras, ou seja, dividindo-o em partes.

Alguns trabalhos já realizados como o de Trombim (2006) e Jesus (2008) utilizando a técnica de amostragem estratificada, demonstraram sucesso em sua realização. Com base nesses dados, esta pesquisa propõe-se a utilizá-la.

Para complementar esta técnica, aplica-se também a amostragem casual ou aleatória simples, que consiste em numerar uma população de **1** a **n** em seguida sortear **k** números desta sequência, resultando nos elementos pertencentes à amostra (CRESPO, 2009).

Os métodos utilizados viabilizam a análise no tráfego de rede na Universidade do Extremo Sul Catarinense, pelo fato da pesquisa ser realizada em dois tipos de redes complexas e de grande porte com imenso volume de informações trafegando pela rede.

6.2 DEFINIÇÃO DE HORÁRIOS E TEMPO DE COLETA

A definição dos horários para coletas de amostragem foi realizada levando-se em consideração as características das redes analisadas.

Testes iniciais foram realizados por meio do software de captura de pacotes e constatou-se que o volume de dados que trafegava era muito grande, gerando um

arquivo com mais de 10 GB, fato que impossibilita o armazenamento das capturas e o tempo para salvamento do arquivo gerado, chegando a 40 minutos em uma das amostras.

6.2.1 Amostragem estratificada e aleatória simples

A partir dos testes iniciais realizados semanalmente, no decorrer do mês, observou-se que o comportamento do tráfego das redes analisadas é heterogêneo. Esta informação foi obtida por meio do software utilizado para monitoramento de ativos utilizado nos servidores da instituição.

Observou-se o comportamento da rede em vários blocos do campus para obtenção daquele que gerasse maior tráfego e que seria escolhido para fazer a análise, a fim de alcançar uma comparação mais próxima do volume total, que por medidas de segurança, não pôde ser realizado no Backbone da instituição.

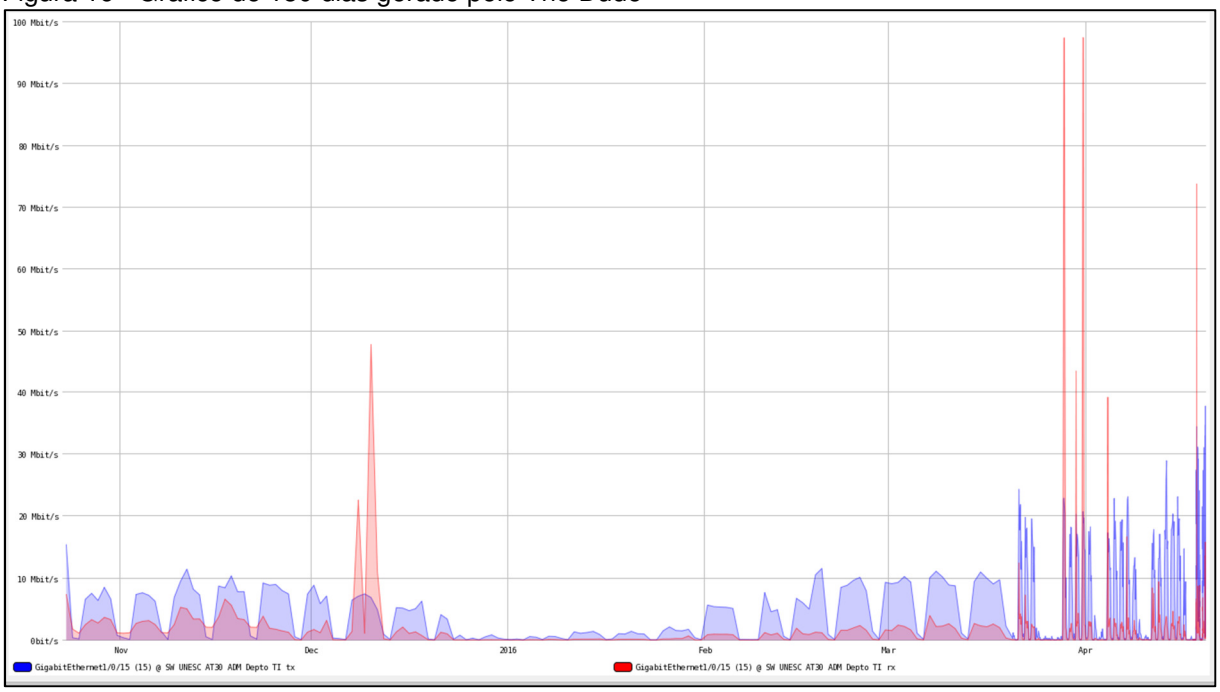
A amostragem estratificada divide a população (X) em subpopulações ou extratos $(X_1, X_2, X_3 \dots X_n)$, em que a variável estudada, de extrato para extrato apresenta um comportamento diferenciado, contudo, dentro desses extratos o comportamento demonstra-se ser mais homogêneo (CRESPO, 2009).

Crespo (2009) explica que a amostragem aleatória simples é semelhante a um sorteio lotérico, numera-se a população de 1 a n e sorteia-se por meio de um dispositivo qualquer.

6.2.2 Horários e amostragem

Para definição do período de coleta considerou-se o tempo disponível, a dificuldade de armazenamento dos dados capturados, as informações obtidas com os testes iniciais e os dados da ferramenta já utilizada pelos analistas de rede para monitoramento de ativos, conforme demonstra Figura 16.

Figura 16 - Gráfico de 180 dias gerado pelo The Dude



Fonte: Microtik.

Para aplicação do método de estratificação, foi utilizado a função Merge da ferramenta Wireshark para unir as amostras testadas inicialmente e aplicar a fórmula p=(E/P).

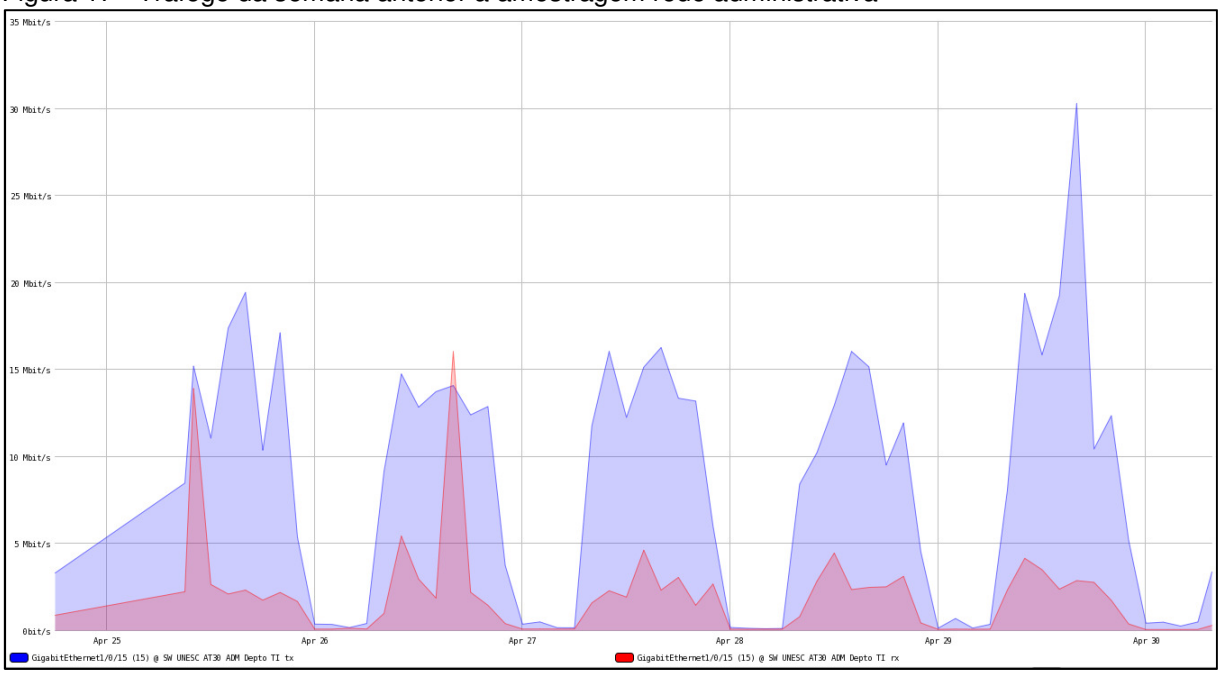
Onde **p** representa a porcentagem do valor do extrato, **E** representa o extrato e **P** a população total.

Obteve-se como resultado que cada coleta teria duração de um minuto, pois o tamanho do arquivo gerado pelo software de captura de dados não fica tão extenso neste tempo, tornando mais rápido o salvamento do arquivo e facilitando o processo de extração e análise dos dados.

A semana definida para realização da amostragem foi do dia 2 a 6 de maio de 2016.

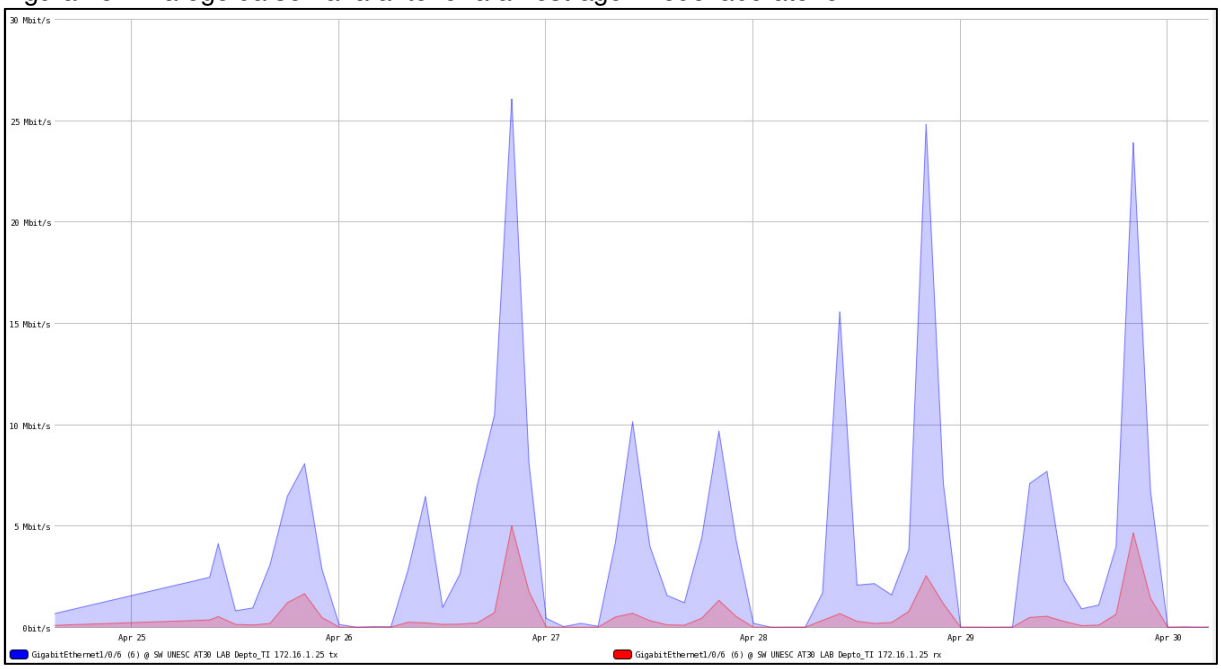
Cada dia foi divido em 3 ciclos considerando-se o período matutino das 8h às 12h, o vespertino das 13h30 às 17h30 e o noturno das 19h às 22h. Esses horários foram definidos conforme gráfico gerado pela ferramenta The Dude, que apresenta o período de maior tráfego nos blocos definidos, conforme Figuras 16 e 17 em que o a cor azul apresenta 2 quedas durante os períodos de maior consumo em todos os dias, estas quedas representam os intervalos.

Figura 17 - Tráfego da semana anterior a amostragem rede administrativa



Fonte: Microtik.

Figura 18 - Tráfego da semana anterior a amostragem rede laboratório



Fonte: Microtik.

Com a utilização da técnica de amostragem estratificada obteve-se os períodos diários, em seguida aplica-se a técnica de amostragem aleatória para definição dos horários para coleta dos dados. Para tal, as horas foram convertidas em

minutos e numeradas de 0 a 240, totalizando 4 horas em cada período definido anteriormente. O comando de sorteio utilizado com o auxílio do softwares Microsoft Office Excel versão 2013 foi “=aleatórioentre(0;240)”.

Com a aplicação da técnica de amostragem aleatória chegou-se ao total de 12 amostras por dia em cada rede, com o tempo de duração para cada coleta de 1 minuto, gerando ao final do período 120 amostras a serem analisadas conforme tabelas a seguir:

Tabela 1 – Período de amostragem da rede administrativa

Bloco S		**Segunda-feira**	**Terça-feira**	**Quarta-feira**	**Quinta-feira**	**Sexta-feira**
Matutino	Coleta 1	8h49min	10h13min	8h08min	10h06min	8h15min
	Coleta 2	9h07min	10h23min	8h37min	10h28min	8h53min
	Coleta 3	9h27min	10h34min	9h14min	11h19min	9h07min
	Coleta 4	9h47min	11h35min	9h51min	11h31min	9h34min
Vespertino	Coleta 1	13h43min	15h34min	14h23min	16h19min	14h01min
	Coleta 2	14h12min	16h07min	14h34min	16h29min	14h12min
	Coleta 3	14h19min	16h29min	15h03min	16h40min	14h24min
	Coleta 4	15h10min	17h11min	15h12min	17h19min	15h08min
Noturno	Coleta 1	19h08min	21h03min	19h25min	21h09min	19h20min
	Coleta 2	19h25min	21h23min	19h36min	21h29min	19h29min
	Coleta 3	19h34min	21h43min	19h50min	21h38min	20h29min
	Coleta 4	20h20min	22h00min	20h25min	21h58min	20h42min

Fonte: do autor.

Tabela 2 - Período de amostragem da rede laboratório

XXI-A		**Segunda-feira**	**Terça-feira**	**Quarta-feira**	**Quinta-feira**	**Sexta-feira**
Matutino	Coleta 1	10h42min	8h05min	10h26min	8h03min	10h09min
	Coleta 2	11h09min	8h12min	10h38min	8h53min	10h42min
	Coleta 3	11h45min	8h45min	11h13min	9h21min	10h52min
	Coleta 4	11h54min	9h19min	11h55min	9h53min	11h25min
Vespertino	Coleta 1	15h58min	14h04min	15h42min	14h07min	15h33min
	Coleta 2	16h51min	14h14min	15h54min	14h17min	16h01min
	Coleta 3	17h03min	14h51min	16h16min	14h46min	16h56min
	Coleta 4	17h17min	14h54min	17h27min	15h22min	17h02min
Noturno	Coleta 1	21h17min	19h26min	21h00min	19h28min	21h12min
	Coleta 2	21h26min	19h44min	21h28min	19h43min	21h24min
	Coleta 3	21h38min	20h08min	21h35min	20h02min	21h33min
	Coleta 4	21h59min	20h23min	21h52min	20h39min	21h54min

Fonte: do autor.

6.3 CENÁRIO

A partir dos horários, períodos e tempo de coleta já definidos, utilizou-se a ferramenta Wireshark em um computador portátil com sistema operacional Microsoft Windows 10, com interface de rede em modo promíscuo na chegada/saída principal de cada bloco escolhido para realização da coleta dos dados para posterior análise.

6.3.1 Estrutura da rede

Atualmente, a rede da UNESC comparada com outras empresas da região, é consideravelmente de grande porte, tanto pela sua estrutura física, quanto ao parque de computadores existentes.

Definiu-se como ponto de coleta da rede laboratório o bloco XXI-A, que é composto por 6 laboratórios de informática com 24 computadores cada um, 15 computadores de salas de aula, e 3 antenas sem fio disponibilizadas para uso dos acadêmicos, que segundo informações fornecidas pelo setor responsável, chega a atingir 100 usuários por antena em períodos de pico.

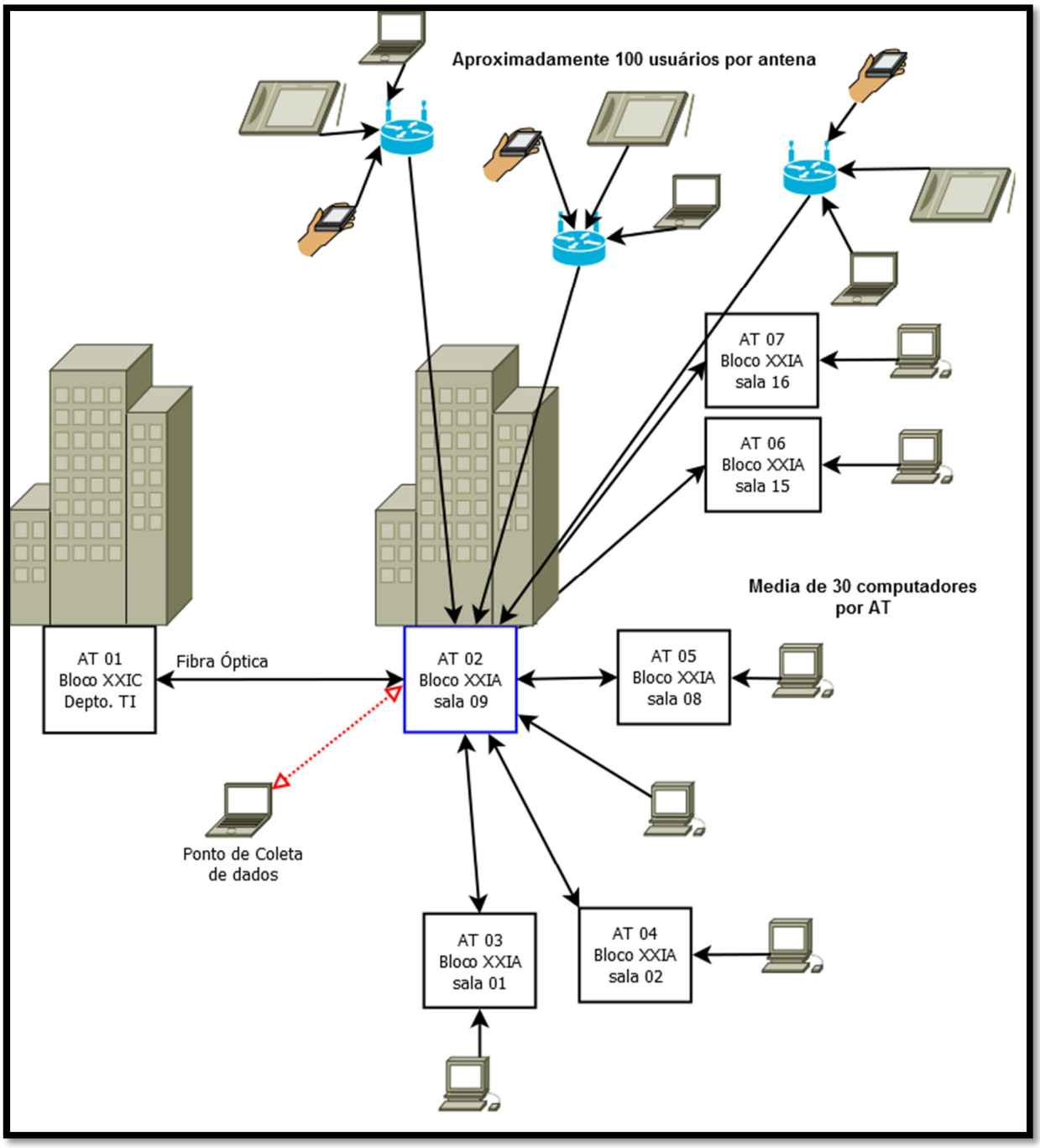
Figura 19 - Número de usuários rede sem fio

:0d:ab:30	BLC_XXIA_01	BLC_XXIA_01	zf7982	Conectado	Disabled	172.16.62.36	172.16.62.36:12223	1	116 (11a/n-40), 1 (11g/n-20) 97
:0d:ce:a0	BLC_XXIA_02	BLC_XXIA_02	zf7982	Conectado	Disabled	172.16.62.34	172.16.62.34:12223	1	124 (11a/n-40), 11 (11g/n-20) 124
:03:7c:e0	BLC_XXIA_03	BLC_XXIA_03	zf7372	Conectado	Disabled	172.16.62.4	172.16.62.4:12223	1	52 (11a/n-40), 4 (11g/n-20) 57

Fonte: Rucks.

Todos os equipamentos de rede e de usuários são interligados em armários de telecomunicação (AT), sendo seis ATs localizados um em cada laboratório, atendendo um lado do corredor e um deles é o principal do bloco que fica interligado via fibra ótica com o Departamento de Tecnologia da Informação, onde foram feitas as capturas para a amostragem da rede laboratório, conforme diagrama a seguir, lembrando que as numerações dos ATs são simbólicas.

Figura 20 - Diagrama da rede XXI-A



Fonte: do autor.

Para realização da coleta de dados neste local, foi utilizado uma função do Switch que permite direcionar todo o tráfego da rede da porta de link, para outra porta livre (Port Mirror), este procedimento tem a finalidade de realização de eventuais testes e até mesmo para a análise do tráfego.

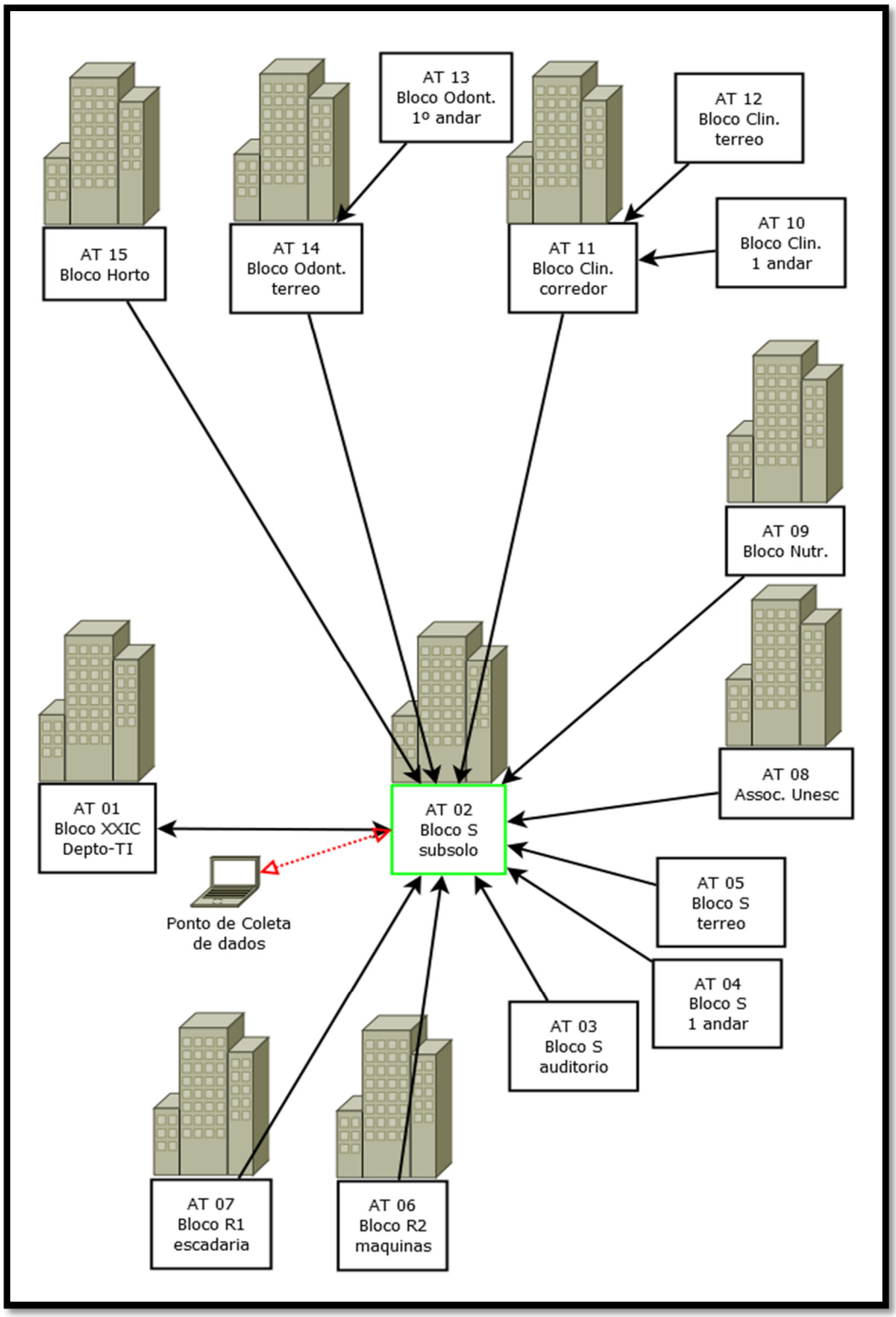
Na rede administrativa, foi utilizado o mesmo recurso de espelhamento de porta no Switch que foi utilizada na rede laboratório. Com resultados obtidos pela ferramenta The Dude, definiu-se que o melhor ponto para realização da coleta de amostragem seria o bloco S.

Este local foi escolhido estrategicamente, por existir o tráfego do bloco S e também de outros blocos atrelados a ele, utilizando o mesmo link de fibra até o departamento de TI.

A rede do bloco S é composta por 3 Armários de Telecomunicações (AT) sendo um em cada andar do bloco que atendem todas as salas de aulas, departamentos e laboratórios de pesquisa.

No subsolo do bloco S, onde fica a distribuição para os blocos vizinhos e também ponto de coleta de dados, estão ligados, bloco R1, bloco R2, clínicas integradas, bloco da odontologia, horto florestal, nutrição, AUNESC, auditório e 3 ATs deste bloco, totalizando 14 armários de telecomunicações, conforme diagrama a seguir onde a numeração dos ATs também são simbólicas.

Figura 21 - Diagrama da rede do bloco S



Fonte: do autor.

A rede do bloco S é composta por aproximadamente 48 departamentos e laboratórios. Ressaltando-se a disponibilização de bancadas para utilização de

computadores pessoais pelos acadêmicos nos laboratórios, com acessos que chegam a passar de 300 computadores simultaneamente em dias de pico.

6.3.2 Aplicação da ferramenta Wireshark

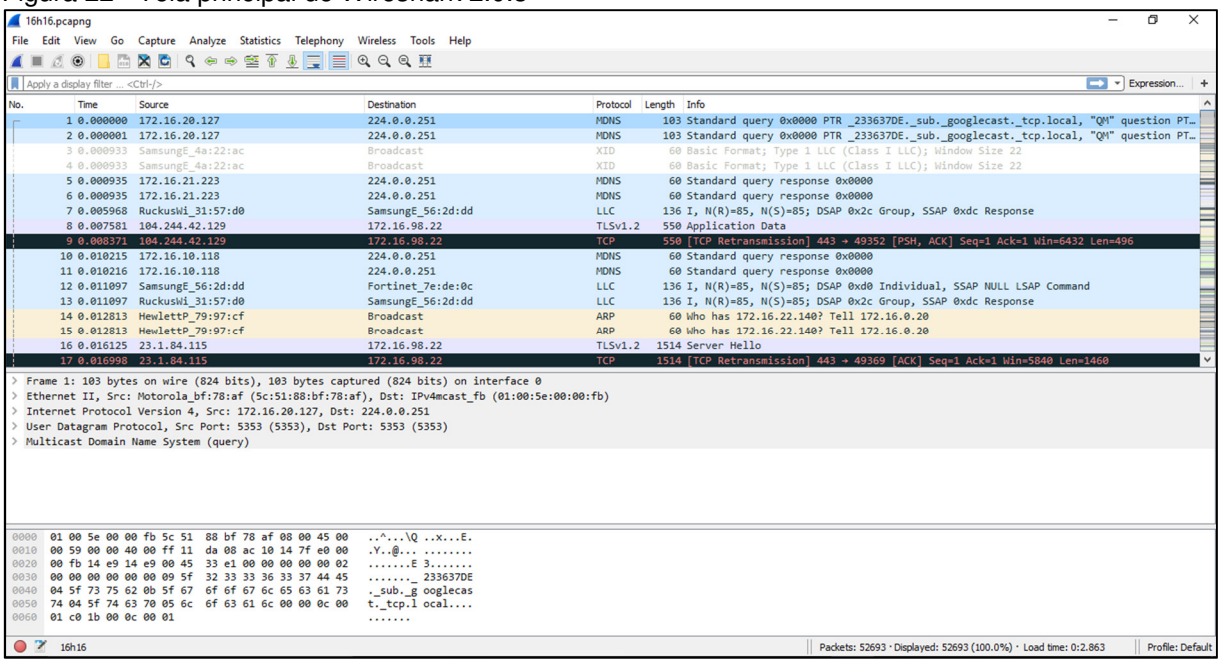
A definição da ferramenta a ser utilizada nessa pesquisa foi especificada a partir de testes realizados com softwares mencionados na seção 3.2 e por referências encontradas na web.

Para desenvolvimento deste trabalho, utilizou-se a ferramenta Wireshark na versão 2.0.3 com a interface de rede do equipamento de coleta em modo promíscuo para realização da captura dos dados. Uma das principais características do Wireshark é sua interface gráfica, que facilita no momento da manipulação dos dados. Ressalta-se que essa ferramenta é *open source* e recebe contribuições de especialistas em rede de todo o mundo, segundo informações da documentação no site oficial.

Destaca-se ainda no *software*, o poder de identificação de vários pacotes que podem ser abertos visualmente, com a possibilidade de acompanhamento de todo o seu processo de encapsulamento.

A Figura 22 demonstra que o primeiro campo possui todos os protocolos capturados e na segunda parte encontra-se a descrição detalhada dos protocolos pertencentes a cada pacote. Para essa visualização, basta selecionar o pacote desejado e na parte inferior da tela principal, observar o conteúdo em hexadecimal de um pacote.

Figura 22 - Tela principal do Wireshark 2.0.3



Fonte: Wireshark.

Uma funcionalidade que foi fundamental para esta pesquisa é a possibilidade de aplicação de filtros na hora da captura, ou mesmo depois que ela foi salva em disco. Esse processo auxiliou na identificação dos serviços utilizados em um tráfego específico, possibilitando o descarte de protocolos que não teriam relevância para a pesquisa, como por exemplo os protocolos POP, SMTP e IMAP que são protocolos de e-mail que mantiveram pouca incidência durante o período de amostragem e que na maioria das vezes não apareceu.

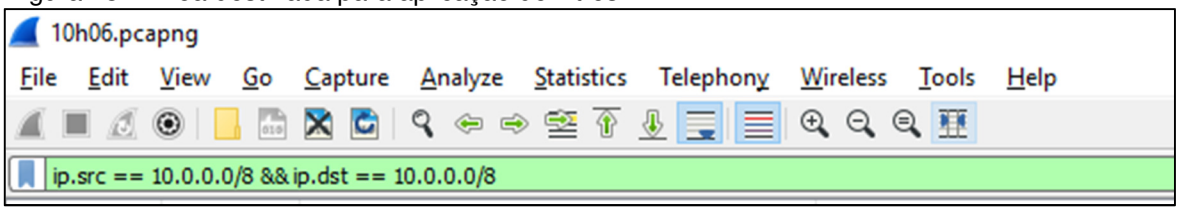
6.3.3 Aplicação de filtros sobre as amostras

A etapa de aplicação de filtros sobre as amostras foi uma das fases mais significativas na execução do projeto, que despendeu maior tempo para sua execução.

Nesta pesquisa optou-se na aplicação dos filtros após os dados já terem sido armazenados em disco, a fim de prevenir a perda de informações por aplicação de filtro de forma errônea.

A utilização deste recurso auxilia na identificação dos protocolos e serviços que estavam sendo utilizados durante o período de amostragem. A Figura 23 apresenta a área onde são inseridos os filtros, conforme a necessidade e o interesse de quem está utilizando a ferramenta.

Figura 23 - Área destinada para aplicação de filtros



Fonte: Wireshark.

Conforme demonstra a Figura 23, observa-se que a utilização do filtro foi de crucial importância para entendimento do comportamento entre as duas redes analisadas. Ele informa todas as requisições e respostas destinadas a rede 10.0.0.0/8 que, neste caso, é a rede administrativa. Apresentando todo o tráfego entre as estações de trabalho, entre servidores, e de estações de trabalhos para os servidores e vice e versa, em suma, todo o tráfego interno da rede administrativa. O mesmo filtro foi utilizado para a rede laboratório.

Na Tabela 3 são listados alguns filtros aplicados para identificação do tráfego nos dados do período de amostragem e nas amostras dos testes iniciais que auxiliaram na definição das escolhas realizadas.

Tabela 3 - Alguns filtros utilizados para identificação do tráfego

Filtro	Descrição do filtro
ip.src == 10.0.0.0/8 && ip.dst == 10.0.0.0/8	Tráfego entre estações e servidores
tcp.port == 80 || udp.port == 80	Navegação Web (WWW)
eth.type != 0x0800	Mostra o tráfego não IP
eth.type == 0x0800	Mostra o tráfego IP
!arp	Mostra todos os pacote, menos os que utilizaram ARP
TCP	Apresenta apenas os pacotes que utilizaram o protocolo TCP

NBSS	Apresenta apenas os protocolos do serviço Netbios Session Service
NBNS	Apresenta apenas os protocolos do serviço Netbios Name Service

Fonte: do autor.

Além dos filtros já apresentados, a ferramenta Wireshark dispõe em seu manual informações e as combinações de todos os filtros, entre outras funcionalidades que a ferramenta possui. Mais detalhes estão descritos na seção 3.2.1.

6.3.4 Armazenamento dos dados

Destaca-se que o armazenamento dos dados foi uma das principais funcionalidades disponibilizada pelo software Wireshark, pois o período de amostragem (descrito na seção 6.2) foi executado com tranquilidade, visto que não houve a realização da análise de cada coleta, executada em uma etapa distinta a esta.

Além da manipulação ter sido realizada em outra etapa da pesquisa, os arquivos ficaram disponíveis para consultas posteriores, em caso de dúvidas ou necessidade de consulta aos arquivos, eles poderiam ser acessados sem dificuldades, pois foram salvos com extensão da própria ferramenta.

No processo de armazenamento, foi gerado um arquivo para cada coleta e cada uma foi nomeada com seu respectivo horário de início de execução, como exemplo 9h53, respeitando rigorosamente o seu início e tempo de duração de um minuto.

Cada arquivo gerado foi armazenado em pastas nomeadas com seu bloco e dia, obtendo no final do período de amostragem 10 (dez) pastas para cada rede e em cada pasta 12(doze) arquivos, tendo um total de 18,6 gigabytes de armazenamento em disco. Ressalta-se que uma das amostras dos testes iniciais teve a duração de aproximadamente 10(dez) minutos, atingiu 3,4 gigabytes e levou em torno de 40 minutos para seu armazenamento. Destaca-se que os métodos descritos na seção 6.2 atendem e viabilizam a realização deste projeto.

6.4 RESULTADOS OBTIDOS

Com os resultados obtidos a partir dos métodos utilizados, observa-se que as duas redes analisadas possuem características bem distintas.

As análises estatísticas foram realizadas por meio do software IBM Statistical Package for the Social Sciencies (SPSS) versão 22.0. As variáveis foram expressas por meio de média, erro padrão, mediana e valores mínimos e máximos. Salienta-se também que alguns gráficos e tabelas foram criados utilizando-se recursos disponibilizados pela ferramenta Wireshark.

Conforme demonstra a Tabela 4, no intuito de se obter o tráfego total da rede administrativa e laboratório, aplicou-se um filtro na ferramenta Wireshark a fim de ignorar os protocolos ARP e RARP que são nativos da rede e que serão esboçados posteriormente.

Tabela 4 - Tráfego total da rede administrativa

Rede Adm. Quantidade de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	300	32406,84[a]	5244,90	2830	16	805527	<0,001
Vespertino	300	29195,02[a]	4364,50	3029,50	21	669414	
Noturno	300	21064,88[b]	3379,20	1149	0	487225	
Volume (MB)							
Matutino	300	24,20[a]	4,65	0,49	0	715,66	<0,001
Vespertino	300	20,49[a]	3,59	0,62	0	616,74	
Noturno	300	14,54[b]	2,60	0,20	0	405,81	

Fonte: do autor.
[a,b,c] –letras distintas representam diferenças estatisticamente significativas após a aplicação do teste Dunn (p<0,05).
* Valor obtido após aplicação do teste H de Kruskal-Wallis.

Na Tabela 4 observa-se que a rede administrativa apresentou uma variação em seu tráfego total, porém, manteve um comportamento padrão no período compreendido das 8h às 17h30, mantendo a média de tráfego entre 20 a 25 megabytes. No período noturno apresentou uma diferença estatisticamente significativa, com tráfego diminuído, chegando a atingir aproximadamente 10 megabytes a menos do que o decorrer do dia (p<0,001).

Outro ponto a ressaltar, foi quando o tráfego alcançou seu pico, atingindo cerca de 11,92 megabytes por segundo. Neste momento, aconteceu um comportamento no switch em que foi feita a captura diferente dos demais períodos de coleta, a retransmissão dos pacotes.

Analisando-se a Tabela 5, visualiza-se que o comportamento da rede laboratório foi o oposto da administrativa, demonstrando uma distinção estatisticamente significativa em seus três períodos durante a semana analisada, chegando a ter uma diferença de 10 megabytes entre eles e apresentando menor tráfego no período vespertino.

Vale ressaltar, que o maior pico de tráfego da rede laboratório ocorreu no período matutino, chegando a atingir cerca de 425 megabytes em apenas um minuto, mantendo a média de aproximadamente 10 megabytes. O maior volume aconteceu no período noturno com média em torno de 21 megabytes por segundo.

Tabela 5 - Tráfego total da rede laboratório

Rede lab. Quantidade de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	300	13734,30[a]	2535,67	1062,50	0	418597	<0,001
Vespertino	300	4307,85[b]	589,57	552	0	88795	
Noturno	300	33003,85[c]	4660,40	3542	0	480700	
Volume (MB)							
Matutino	300	10,40[a]	2,49	0,15	0	424,64	<0,001
Vespertino	300	1,83[b]	0,43	0,09	0	80,65	
Noturno	300	20,61[c]	3,30	0,51	0	342,73	

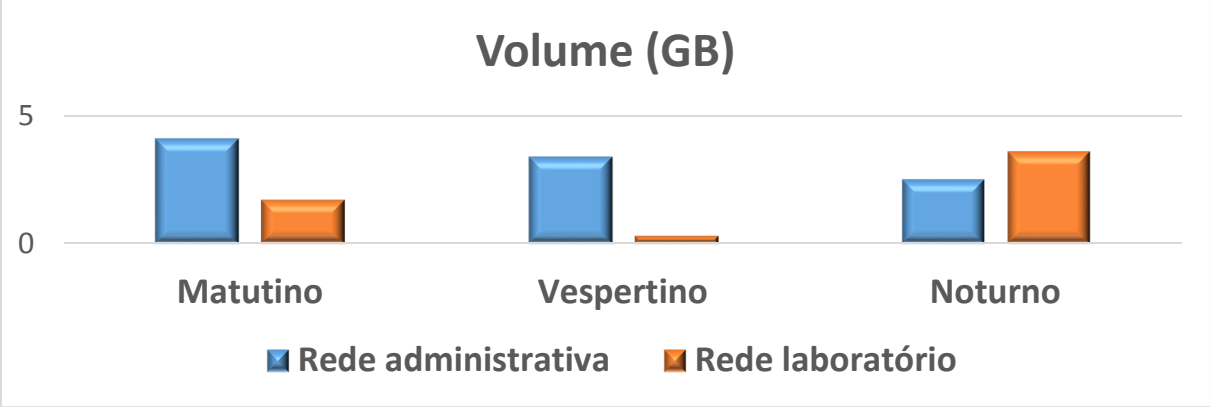
Fonte: do autor.
[a,b,c] –letras distintas representam diferenças estatisticamente significativas após a aplicação do teste Dunn (p<0,05).
* Valor obtido após aplicação do teste H de Kruskal-Wallis.

Nos gráficos 1 e 2 são apresentados um resumo da quantidade de pacotes capturados nas duas redes e o volume gerado no período de amostragem. Foram 14.020.501 pacotes capturados na rede administrativa que gerou um volume de 9,98 gigabytes, e na rede laboratório foram 9.284.551 pacotes capturados e um volume de 5,61 gigabytes.

Gráfico 1 – Quantidade de protocolos capturados

Fonte: do autor.

Gráfico 2 – Volume total capturado



Fonte: do autor.

A Tabela 6, indica que o tráfego de navegação WWW não apresentou uma diferença estatisticamente significativa ($p<0,05$). Nas médias apresentadas na tabela, estão o protocolo HTTP (transporte), e o TCP e UDP com destino a porta 80 (serviço World Wide Web).

Tabela 6 - Tráfego de Navegação (WWW) bloco S

Rede Adm. Quantidade de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	20	5712,69	1010,2	4315,38	1858	20273	<0,181
Vespertino	20	5606,28	973,27	4580,3	1831	22443	
Noturno	20	4100,68	759,54	3323,52	310	12594	
Volume (MB)							
Matutino	20	4,36	0,95	3,35	1,03	20,75	<0,172
Vespertino	20	4,46	0,94	3,12	1,05	20,72	
Noturno	20	3,04	0,63	2,4	0,19	11,3	

Fonte: do autor.
* Valor obtido após aplicação do teste H de Kruskal-Wallis.

Conforme Tabela 7, os dados da navegação HTTPS mantiveram um comportamento padrão. Neste tráfego foram identificados os protocolos TCP com destino a porta 443, os SSL/TLS que são de segurança e que protegem a navegação por página e os serviços de e-mail.

Tabela 7 - Tráfego de navegação HTTPS

Rede Adm. Quantidade de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	20	6805,06	1235,39	5118,91	2173	24591	<0,185
Vespertino	20	6636,57	1189,63	5326,28	2062	27289	
Noturno	20	4894,39	921,15	3974,41	341	15227	
Volume (MB)							
Matutino	20	10,00	1,96	7,50	2,23	35,07	<0,088
Vespertino	20	8,25	1,28	7,05	2,72	30,22	
Noturno	20	6,12	0,96	5,24	0,82	19,88	

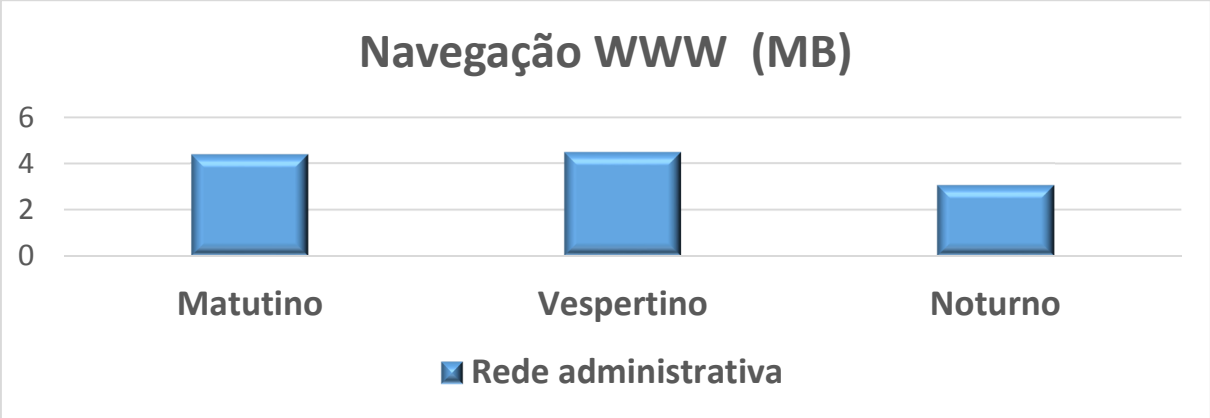
Fonte: do autor.
* Valor obtido após aplicação do teste H de Kruskal-Wallis.

Com a utilização de alguns recursos disponibilizados pela ferramenta Wireshark, observou-se que em horário de maior tráfego a navegação web do bloco

S, representa em torno de 4% do tráfego total do ponto analisado e a HTTPS aproximadamente 4,9%.

No gráfico 3 deixa mais claro o comportamento do tráfego gerado pela navegação web na rede administrativa.

Gráfico 3 – resumo navegação web rede administrativa



Fonte: do autor.

Identificou-se algumas URLs dentro do tráfego de navegação web que poderiam caracterizar um mau uso do recurso, sendo que algumas URLs são de mídias sociais e sites de notícia, conforme apresenta a Tabela 8.

Tabela 8 - Sites mais acessados no período de coleta

www.facebook.com
www.gmail.com
www.youtube.com
www.instagram.com
play.google.com
mail.google.com
www.outlook.com
www.terra.com.br
g1.globo.com
www.engeplus.com.br

Fonte: do autor.

Na Tabela 9 apresenta-se o tráfego interno gerado pelo bloco S. Identificou-se com maior incidência os protocolos de compartilhamento de arquivos SMB, de transporte TCP e UDP, o serviço de DNS, IP e o de gerência de rede SNMP, apontado como o terceiro com maior ocorrência na rede.

Constatou-se que no bloco S cerca de 86% do tráfego foi interno, identificando acesso aos servidores de arquivos, sistema acadêmico entre outros serviços.

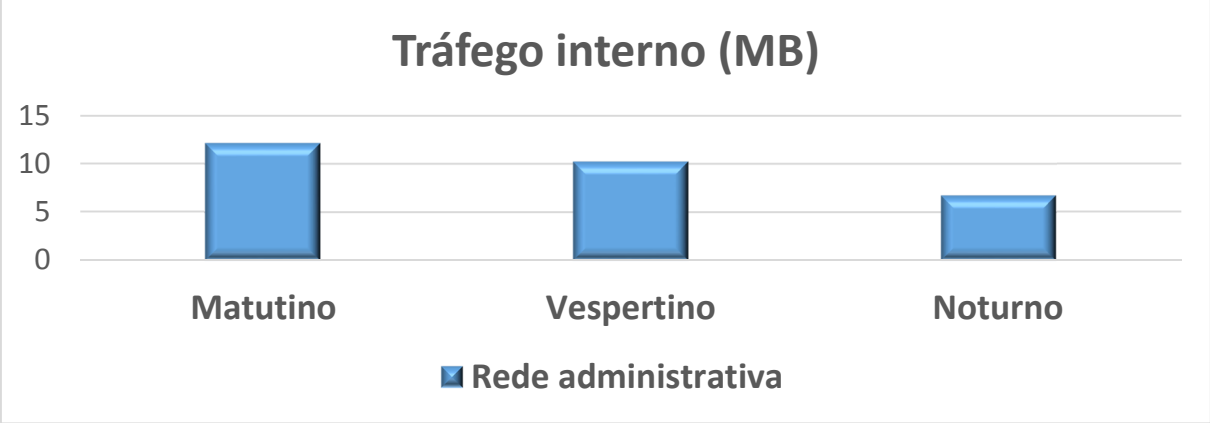
Tabela 9 - Resumo do tráfego interno bloco S

Quantidade de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	20	15907,73	3208,61	11021,67	6077	63735	<0,067
Vespertino	20	14007,31	2223,20	11806,12	5730	52874	
Noturno	20	9929,97	1715,79	7736,86	997	30554	
Volume (MB)							
Matutino	20	12,16	3,26	7,46	2,46	57,98	<0,113
Vespertino	20	10,26	2,25	7,57	2,59	49,99	
Noturno	20	6,75	1,40	5,31	4,45	25,88	

Fonte: do autor.
* Valor obtido após aplicação do teste H de Kruskal-Wallis.

No gráfico 4 mostra o resumo do tráfego interno gerado pela rede administrativa.

Gráfico 4 – resumo trafego interno rede administrativa



Fonte: do autor.

As tabelas 10, 11 e 12 demonstradas a seguir, exibem a quantidade de pacotes capturados com alguns protocolos e serviços identificados no bloco S, apontando as maiores incidências.

Tabela 10 - Quantidade de protocolos capturados período matutino

Protocolo	Total	Média	Erro padrão	Mediana	Mínimo	Máximo
TCP	2777577	138878,85	25211,94	104467,50	44348	501863
IP	6457	322,85	88,90	131,00	22	1361
HTTP	58493	2924,65	192,30	3070,00	1272	4959
SMB	161874	8093,70	1915,10	5397,50	1526	27915
SSL/TLS	331371	16568,55	1224,30	16981,00	6297	29161
SNMP	203973	10198,65	1248,17	8646,50	3519	20540
UDP	158257	7912,85	3138,58	3075,50	2234	50021
DNS	43300	2165,00	126,59	2135,50	1151	3420
ICMP	13817	690,85	54,04	690,50	341	1165
IGMP	1885	94,25	14,48	77,00	16	290
ARP	198624	9931,20	213,12	9915,00	8515	12025
NBNS	80950	4047,50	1410,77	2662,50	960	30259
NBSS	356648	17832,40	16137,16	34,00	21	322826
IPX	4620	231,00	7,95	231,00	151	306

Fonte: do autor.

Tabela 11 - Quantidade de protocolos capturados período vespertino

Protocolo	Total	Média	Erro padrão	Mediana	Mínimo	Máximo
TCP	2708803	135440,15	24278,137	108699,50	42077	556925
IP	12860	643,00	188,319	392,00	21	3358
HTTP	60963	3048,15	242,621	2849,00	1774	6102
SMB	217222	10861,10	3205,755	3244,00	1516	54118
SSL/TLS	362770	18138,50	1380,833	18253,50	8766	32246
SNMP	181764	9088,20	1037,186	8906,00	2963	20256
UDP	114976	5748,80	1304,517	4247,50	2635	28241
DNS	48312	2415,60	132,404	2315,00	1619	3797
ICMP	20514	1025,70	91,462	911,50	417	1948
IGMP	2574	128,70	14,104	123,00	38	246
ARP	238905	11945,25	557,733	11919,50	6092	19550
NBNS	65480	3274,00	241,760	3401,50	1674	5604
NBSS	43842	2192,10	2097,701	45,50	36	42040
IPX	5560	278,00	8,185	286,00	222	336

Fonte: do autor.

Tabela 12 - Quantidade de protocolos capturados período noturno

Protocolo	Total	Média	Erro padrão	Mediana	Mínimo	Máximo
TCP	1997709	99885,45	18798,990	81110,50	6957	310745
IP	3008	150,40	46,418	75,00	22	908
HTTP	39382	1969,10	283,927	1645,50	176	5145
SMB	47904	2395,20	1498,411	854,00	188	30779
SSL/TLS	372557	18627,85	5434,172	9887,00	1092	113515
SNMP	149473	7473,65	1008,925	6909,50	442	17277
UDP	66146	3307,30	1074,321	1707,50	979	22989
DNS	24628	1231,40	181,313	987,50	196	3255
ICMP	10243	512,15	55,138	439,00	207	1134
IGMP	1320	66,00	12,218	54,00	0	198
ARP	166591	8329,55	258,727	8363,00	6160	10958
NBNS	18345	917,25	68,791	948,00	338	1538
NBSS	361	18,05	4,095	15,00	3	92
IPX	3583	179,15	8,556	188,50	84	227

Fonte: do autor.

Analisando as informações contidas nas tabelas 10, 11 e 12, observa-se que o tráfego da rede administrativa utiliza a arquitetura TCP/IP em 63% no período matutino, 66% no vespertino e 69% no noturno, em contrapartida ao protocolo UDP, que usa 4% no período matutino, 3% no vespertino e 2% no noturno.

Prosseguindo com a análise das tabelas 10, 11 e 12, destacam-se os protocolos SSL/TLS, que apresentaram o maior número de pacotes capturados no período de amostragem com um consumo médio entre 15 MB e 17 MB, sendo que no período noturno a sua incidência diminuiu. Outro detalhe a observar é que o terceiro protocolo (SNMP) com maior número de pacotes, representou cerca de 1% do volume na rede. Além disso, 10% do tráfego são consumidos por protocolos não IP, evidenciando o ARP.

A partir dos dados expostos até o momento, torna-se evidente que o maior consumo da rede no bloco S está relacionado aos acessos internos, que de certa forma seria um tráfego produtivo por serem destinados a servidores de dados, sistema acadêmico, entre outros. Porém, o que ficou evidente durante o período de coleta, foi o comportamento anormal da ferramenta de captura de dados ao identificar retransmissões de pacotes do Switch quando o link entre o AT principal e o bloco S estava perto do seu limite.

Prosseguindo-se com as análises, as tabelas 13 e 14, apresentam o tráfego de navegação web e HTTPS dos laboratórios e também outros resultados do bloco XXI-A.

Tabela 13 - Tráfego de navegação (WWW) XXI-A

Quantidade de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	20	18865,76[a]	5291,2	10416,7	2378	91947	<0,001
Vespertino	20	4116,18[b]	971,26	2682,56	354	20137	
Noturno	20	41475,08[c]	5473,3	37772,9	4512	88042	
Volume (MB)							
Matutino	20	17,10[a]	5,74	7,13	0,76	94,42	<0,001
Vespertino	20	2,57[b]	1,01	0,88	0,03	20,36	
Noturno	20	29,48[a]	4,34	28,82	0,82	71,42	

Fonte: do autor.
[a,b] –letras distintas representam diferenças estatisticamente significativas após a aplicação do teste Dunn (p<0,05).
* Valor obtido após aplicação do teste H de Kruskal-Wallis.

Tabela 14 - Tráfego de navegação HTTPS XXI-A

Quantidade de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	20	3647,48[a]	1078,65	1963,99	333	18538	<0,001
Vespertino	20	632,84[b]	198,07	331,28	11	3961	
Noturno	20	7698,60[c]	1050,88	7151,13	723	16871	
Volume (MB)							
Matutino	20	4,89[a]	1,36	2,40	0,60	24,54	<0,001
Vespertino	20	0,88[b]	0,23	0,52	0,15	0,47	
Noturno	20	10,24[c]	1,23	10,45	1,16	19,81	

Fonte: do autor.
[a,b,c] –letras distintas representam diferenças estatisticamente significativas após a aplicação do teste Dunn (p<0,05).
* Valor obtido após aplicação do teste H de Kruskal-Wallis.

No gráfico 5 deixa mais claro o comportamento do tráfego gerado pela navegação web na rede laboratório.

Gráfico 5 – resumo navegação web rede laboratório

Fonte: do autor.

Após aplicação dos testes estatísticos de Dunn e Kruskal-Wallis, observou-se que tanto em navegação WWW como na HTTPS (Tabela 14), existiu uma diferença estatisticamente significativa no número de pacotes capturados nos três períodos de análise. O volume não se manteve em um comportamento padrão entre os ciclos, exibindo tráfego maior no período noturno e menor no vespertino.

Nas amostragens coletadas, identificou-se como principais protocolos o TCP direcionado para as portas 80 e 443; o UDP destinado a porta 80; o de transporte HTTP; os de segurança SSL/TLS e além dos já mencionados, o eXtensible Messaging and Presence Protocol (XMPP) definido pela Request for Comments (RFC) 3920 como protocolo aberto, que tem a finalidade de construir aplicativos de mensagens instantâneas. Este, não teve tanta incidência na rede administrativa, mas destacou-se na de laboratório.

Utilizando-se os recursos disponibilizados pela ferramenta Wireshark, identificou-se que na captura que apresentou maior tráfego a navegação web representou 28,3% e HTTPS representou 57,8% do tráfego, evidenciando que o maior tráfego da rede laboratório ocorre em acesso à Internet.

Na Tabela 15 são listadas algumas URLs identificadas no período de coleta do bloco XXI-A.

Tabela 15 - Sites mais acessados no período de coleta XXI-A

www.unesc.net
www.instagram.com
www.youtube.com
www.ahnegão.com.br
www.naointendo.com.br
www.umsabadoqualquer.com
www.vidadeprogramador.com.br
mail.google.com
g1.globo.com
www.engeplus.com.br

Fonte: do autor.

Realizou-se um filtro na ferramenta Wireshark para identificação do tráfego interno da rede laboratório, que é apresentado pela Tabela 16. Nesta análise aponta-se a presença dos serviços de DNS e DHCP, o protocolo TCP, o de gerência de rede SNMP e um outro que é integrante do IP definido pela RFC-792 (ICMP), que serve para fornecer relatórios de erros.

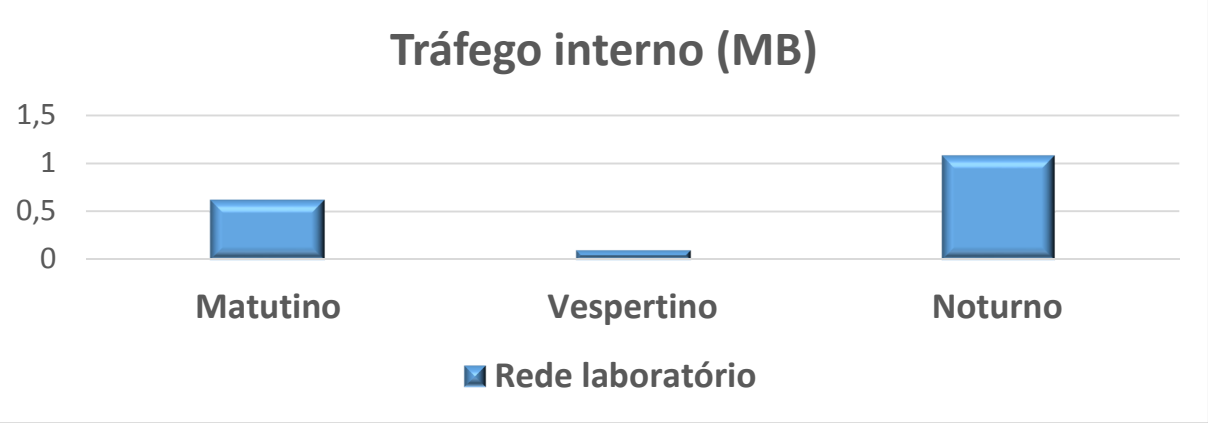
Tabela 16 - Tráfego interno XXI-A

Quantidade de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	20	748,57[a]	190,24	433,69	111	3406	<0,001
Vespertino	20	185,24[b]	35,86	138,12	27	743	
Noturno	20	1718,63[c]	215,45	1542,21	205	3435	
Volume (MB)							
Matutino	20	0,62[a]	0,20	0,26	0,03	3,37	<0,001
Vespertino	20	0,10[b]	0,36	0,04	0	0,72	
Noturno	20	1,08[a]	0,16	1,06	0,04	2,56	

Fonte: do autor.
[a,b,c] –letras distintas representam diferenças estatisticamente significativas após a aplicação do teste Dunn (p<0,05).
* Valor obtido após aplicação do teste H de Kruskal-Wallis.

No gráfico 6 mostra o resumo do tráfego interno gerado pela rede laboratório.

Gráfico 6 – resumo tráfego interno rede laboratório



Fonte: do autor.

Observa-se na Tabela 16 que o comportamento do tráfego interno apresenta diferença entre os turnos, concentrando o maior número de pacotes capturados no período noturno. Em comparação com a rede do bloco S salienta-se que não teve tanta representatividade tendo gerado um tráfego de aproximadamente 1% do total, em que identificou-se acessos ao ambiente virtual, diário on-line e o monitoramento de ativos com o protocolo SNMP.

Nas próximas tabelas exibe-se a quantidade de pacotes capturados com alguns protocolos e serviços que são utilizados na rede laboratório. Optou-se pela apresentação dos mesmos protocolos do bloco S, a fim de destacar a diferença e comportamento entre as redes estudadas.

Tabela 17 - Quantidade de protocolos capturados período matutino

Protocolo	Matutino	Média	Erro padrão	Mediana	Mínimo	Máximo
TCP	1262105	63105,25	18661,773	33979,00	6468	320737
IP	15365	768,25	182,918	548,00	0	2548
HTTP	11582	579,10	149,893	452,50	60	3154
SMB	9490	474,50	270,125	76,50	0	4717
SSL/TLS	151584	7579,20	2490,922	5376,50	725	53007
SNMP	27806	1390,30	326,319	244,50	209	3572
UDP	85758	4287,90	942,709	2627,50	36	14853
DNS	13396	669,8	53,743	651,50	387	1201
ICMP	44217	2210,85	188,571	2111,50	932	3795
IGMP	7716	385,80	30,491	363,50	204	861
ARP	302403	15120,15	866,007	15430,00	8919	21817
NBNS	36790	1839,50	142,377	2034,00	450	2661
NBSS	123	6,15	2,019	2,00	0	30

IPX	7168	358,40	44,703	331,00	125	870

Fonte: do autor.

Tabela 18 - Quantidade de protocolos capturados período vespertino

Protocolo	Vespertino	Média	Erro padrão	Mediana	Mínimo	Máximo
TCP	218976	10948,80	3426,777	5731,50	197	68533
IP	25321	1266,05	300,294	954,50	0	5009
HTTP	3856	192,80	37,821	175,00	0	583
SMB	1256	62,80	11,470	53,00	2	179
SSL/TLS	41353	2067,65	422,701	2062,50	129	6791
SNMP	20262	1013,10	291,338	245,00	0	3646
UDP	25425	1271,25	340,517	621,50	175	5411
DNS	8114	405,70	62,725	379,50	5	981
ICMP	42743	2137,15	114,788	2072,00	1053	3260
IGMP	6385	319,25	24,067	305,50	51	540
ARP	242177	12108,85	804,331	12239,50	3752	16510
NBNS	25732	1286,60	102,731	1130,50	808	2345
NBSS	60	3,00	1,110	0,00	0	16
IPX	2320	116,00	19,467	76,00	0	299

Fonte: do autor

Tabela 19 - Quantidade de protocolos capturados período noturno

Protocolo	Noturno	Média	Erro padrão	Mediana	Mínimo	Máximo
TCP	2663897	133194,85	18183,059	123722	12515	291900
IP	131807	6590,35	1278,214	6987	2	21679
HTTP	48519	2425,95	251,376	2338,5	461	4207
SMB	7739	386,95	80,742	320	103	1797
SSL/TLS	459000	22950	3082,351	20382	5175	65361
SNMP	36334	1816,7	353,549	1736,5	208	4369
UDP	265781	13289,05	2347,771	12713,5	546	40617
DNS	79389	3969,45	537,207	3310	1420	10507
ICMP	86878	4343,9	364,504	3653	2542	8095
IGMP	14361	718,05	43,212	674,5	403	1155
ARP	613879	30693,95	3207,877	26533	672	57388
NBNS	124433	6221,65	528,032	6663	1679	9044
NBSS	225	11,25	2,151	12	0	27
IPX	15984	799,2	70,219	719	313	1534

Fonte: do autor.

Com base nas informações das tabelas 17, 18 e 19 fica evidente que assim como a rede administrativa, a de laboratório também utiliza a arquitetura TCP/IP para

seu funcionamento. Na linha 1 o uso do TCP apresenta 64% no período matutino, 33% no vespertino e 59% no noturno. Em contrapartida, o UDP, apresentado na linha 7 das tabelas, com 4% no período matutino, 4% no vespertino e 6% no noturno.

Identificou-se tráfego não IP de 21,67%, constatando-se uma grande incidência de utilização do protocolo ARP e RARP que representa 2,67% do tráfego, aproximadamente 3,32 megabytes.

O terceiro protocolo com maior incidência no bloco XXI-A foi o de segurança SSL/TLS que representou 14,67% do tráfego, aproximadamente 32,30 megabytes.

A partir dos dados obtidos com a análise da rede laboratório, evidencia-se que seu uso maior é em navegação web, concentrando a maior parte do tráfego no período noturno.

7 CONCLUSÃO

Atualmente, a rede da UNESC comparada com outras empresas da região é consideravelmente de grande porte, tanto pela sua estrutura física, quanto ao parque de computadores existentes. Por meio de um *software* para monitoramento de ativos já utilizado pela instituição em seus servidores, observou-se que o comportamento de rede é heterogêneo durante o dia. Isto se deve ao fato da quantidade de usuários existentes e aos diversos perfis que utilizam o recurso. Foi a partir dessas informações que surgiu a ideia da realização desta pesquisa, que visou o estudo do tráfego e o comportamento das redes administrativa e de laboratórios existentes na instituição.

No momento em que se desenvolve uma pesquisa num ambiente que está em atividade constante, são várias as dificuldades encontradas.

Um dos primeiros problemas encontrados foi ao receber a informação de que não seria possível realizar a coleta dos dados no Armário de Telecomunicação principal da instituição, por políticas de segurança e por se tratar de um ambiente de produção. Nesse momento, surgiu a necessidade de estudar melhor toda a estrutura da rede e definir um novo ponto para fazer a coleta das informações. De certa forma este problema inicial foi resolvido, porém, desencadeou outros que são mencionados adiante.

Após a definição dos novos locais para análise, outra dificuldade surgiu, a forma de como deveria ser feito a coleta por ter um volume muito alto de dados (informações obtidas de algumas coletas que serviram para definição dos novos pontos). Com o auxílio de profissionais da área estatística, métodos foram definidos para que melhor se adequassem a situação e que validariam a pesquisa chegando o mais próximo do tráfego total possível.

O passo seguinte foi o estudo de uma alternativa para colocar a interface de rede do equipamento de coleta em modo promíscuo, para captura do tráfego geral nos pontos de coleta definidos, sendo que a rede da universidade é composta por *Switchs*, e com isso, disponibilizaria apenas o tráfego direcionado a este equipamento.

Depois de muito estudo e com auxílio dos profissionais da área de redes da instituição, foi encontrada a solução: como nos pontos de coletas os *Switchs* são

gerenciáveis, foi realizado o direcionamento do tráfego da porta em que o link principal é ligado para alguma outra porta livre do próprio *Switch*, técnica conhecida como *Port mirror*. Contudo, como essa técnica nunca tinha sido utilizada, não havia conhecimento do que poderia ocorrer se aplicado com a rede em funcionamento.

Vários testes foram realizados em bancada, a fim de verificar problemas que pudessem ocorrer. Posteriormente, constatou-se que não teve alterações no comportamento do *Switch*, tanto no processamento como na memória. Então, foi definido um horário de menor pico na instituição para realização de um teste, que caso não resultasse positivamente, inviabilizaria a pesquisa.

Superadas as dificuldades apresentadas, foi realizada a coleta dos dados e por fim, a análise dos resultados obtidos, que possibilitaram o entendimento do comportamento das redes analisadas e o que trafega por elas.

Destacando-se que o maior consumo da rede no bloco S está relacionado aos acessos internos e que o switch teve que realizar retransmissões de pacotes quando o tráfego estava perto do seu limite, aponta-se como opção de melhoria o aumento do link para o AT principal da instituição, que atualmente é de 100Mb/s. Outra proposta está relacionada a conscientização dos usuários para acesso à Internet, que tanto pode estar sendo usada para o trabalho, como para distrações, já que as coletas aconteceram em períodos de produção, ignorando os intervalos intrajornada.

A partir dos dados obtidos com a análise da rede laboratório, evidencia-se que seu uso maior é em navegação web no período noturno. Neste sentido, como sugestão de melhoria, cita-se em primeiro momento, uma conscientização para os usuários sobre o uso da Internet, pois analisando-se as URLs mais acessadas, fica evidente o mau uso do recurso segundo resolução número 14/2011/CONSU que institui a política de uso dos recursos computacionais e segurança da informação da UNESC, e está disponível para consulta em seu portal http://www.unesc.net/portal/resources/official_documents/6168.pdf?1322135241.
Além disso, identifica-se tráfego direcionado a mensageiros instantâneos tendo como principal mensageiro o Whatsapp e Snapchat. Visualizou-se alto índice de streaming de vídeo, com principal uso pelo Youtube e alguns sites de notícias, além de downloads de diversos tipos de mídia. Outra sugestão seria o estudo de uma forma de união do link de Internet, já que na rede administrativa, o uso da Internet tende a

diminuir no período noturno e na de laboratório aumenta. Outra forma, poderia ser a verificação de um modo de priorizar os serviços que tenham maior importância e consequentemente, limitar o que não for produtivo.

Dentre estas sugestões, ressalta-se ainda que a substituição de equipamentos antigos por mais modernos e gerenciáveis e o fato de mantê-los com seu firmware atualizado, impacta de forma positiva na velocidade e desempenho em ambas as redes mencionadas.

Existem vários estudos na comunidade científica abordando esse assunto e trabalhos já realizados que possuem semelhança com a presente pesquisa, contudo, aponta-se como diferença mais significativa em relação a essa, o volume de dados, a quantidade de equipamentos e também o maior número de amostragens obtidas, pois trata-se da análise de duas redes, a de laboratórios e uma administrativa. É válido ressaltar, que em relação as pesquisas já realizadas sobre o assunto abordado, foi acrescentada a análise de um outro ponto da rede da instituição, além da que já tinha sido estudada e a utilização de métodos estatísticos para apresentação dos dados, possibilitando um diagnóstico completo de toda a rede da Universidade. Esta pesquisa demonstra de forma resumida, aos profissionais da área, o comportamento das duas redes existentes.

Evidencia-se que os objetivos propostos para esta pesquisa foram alcançados, adquirindo-se conhecimento sobre os protocolos de transmissão de dados, a ferramenta para a captura dos protocolos e dados (Wireshark), os métodos estatísticos para uma amostragem válida e os testes estatísticos para demonstração dos resultados e diagnóstico do comportamento da rede.

Com os dados obtidos, destaca-se o comportamento distinto entre as duas redes, mas que possuem as mesmas características na sua arquitetura e funcionalidade.

Como sugestões para trabalhos futuros, cita-se:

a) aplicar outros métodos estatísticos tanto para a amostragem, como para apresentação dos dados, diferentes dos apresentados nesta pesquisa;
b) utilizar outras ferramentas para análise dos pacotes ou monitoramento de tráfego;

c) efetuar uma amostragem por tempo maior, a fim de diminuir o desvio padrão, chegando a uma média amostral mais representativa.

REFERÊNCIAS

ARANGO, Héctor Gustavo. **Bioestatística:** teoria e computacional. 2 ed. Rio de Janeiro: Guanabara Koogan, 2005. 423 p.

BARBETTA, Pedro Alberto; REIS, Marcelo Menezes; BORNIA, Antonio Cezar. **Estatística:** para cursos de engenharia e informática. 3. ed São Paulo: Atlas, 2010. 410 p.

BONOMO , Esley. **Gerenciamento e monitoração de redes de computadores utilizando-se zabbix.** 2006. 62 f. Monografia (Lato Sensu em Administração de Redes Linux) Universidade Federal de Lavras, 2006.

CALLEGARI-JACQUES, Sidia M. **Bioestatística:** princípios e aplicações. Porto Alegre: Artmed, 2004. 255 p.

CARISSIMI, Alexandre da Silva; ROCHOL, Juergen; GRANVILLE, Lisandro Zambenedetti. **Redes de computadores.** 2. ed. Porto Alegre: Bookman, 2009. 391p.

CASAGRANDE, Rogério A. **Técnicas de detecção de sniffers.** Dissertação (Mestrado em Ciência da Computação). Programa de Pós-Graduação em computação, Instituto de Informática, Universidade Federal do Rio Grande do Sul. Porto Alegre, 2003.

CASE, J.; Fedor, M.; Schoffstall, M.; Davin, J. **A simple network management protocol SNMP**. Request for Comments RFC 1157.

CHOWDHURY, Dhiman Deb. **Projetos avançados de rede IP:** roteamento, qualidade de serviço e voz sobre IP. Rio de Janeiro: Campus, 2002. 380 p.

COCHARAN, William G. **Sampling Techniques.** 3. ed. New York: John Willey, 1977.

COMER, Douglas. **Interligação em rede com TCP/IP.** Rio de janeiro: Elsevier, 2006.

_____. **Interligação de rede com TCP/IP.** 6 ed. Rio de janeiro: Elsevier, 2015. 520 p.

COMER, Douglas; STEVENS, David L. **Interligação em rede com TCP/IP.** Rio de janeiro: Campus, 1999.

COMER, D. E. **Internetworking with TCP/IP.** 4. ed. New Jersey: Prentice Hall, 2000. v. 1. 750 p.

CRESPO, Antônio Arnot. **Estatística fácil.** 19. ed. São Paulo: Saraiva, 2009. 218 p.

DEVORE, Jay L. **Probabilidade e estatística:** para engenharia e ciências. São Paulo: Thomson, 2006. 692 p.

DIMARZIO, J. F. **Projeto e arquitetura de redes:** um guia de campo para profissionais de TI. Rio de Janeiro: Campus, 2001. 370 p.

FARREL, Adrian. **A Internet e seus protocolos:** uma análise comparativa. Campus, 2005. 608 p.
FIELD, Andy. **Descobrindo a estatística usando o SPSS.** 2.ed. Porto Alegre: Artmed, 2009. 688 p.

FOROUZAN, Behrouz A. **Comunicação de dados e redes de computadores.** 3 ed. Porto Alegre: Bookman, 2006. 840p.

FRANCESCHI, André de A. **Um modelo de tráfego de rede para aplicação de técnicas de controle estatístico de processos.** Tese (Doutorado em Física). Instituto de Física de São Carlos, Universidade de São Paulo. São Paulo, 2003.

FURMANKIEWICZ, Edson. **Segurança máxima:** o guia de um hacker para proteger seu site na Internet e sua rede. Rio de Janeiro: Campus, 2000. 826 p.

GALLO, Michael A. **Comunicação entre computadores e tecnologias de rede.** São Paulo: Pioneira Thonson Learning, 2003. 673 p.

HAINAUT, L. d'. **Conceitos e métodos da estatística:** uma varíável a uma dimensão. 2. ed Lisboa: Fundação Calouste Gulbenkian, 1997. 362p.

HAYDEN, Matt. **Aprenda em 24 horas redes.** 2.ed Rio de Janeiro: Campus, 1999. 461 p.

JESUS, Fabricio Cardoso de. **Análise da rede sob o ponto de vista do controle de informações e tráfego estudo de caso:** TSA Química do Brasil. 2008. 85 f. Trabalho de Conclusão de Curso (Graduação) -
Universidade do Extremo Sul Catarinense - Unesc, Criciúma, 2008.

JUNIOR, Ademar de Souza Reis; FILHO, Milton Soares. **Um sistema de testes para a detecção remota de Sniffers em redes tcp/ip.** 2002. 68f. Monografia (Graduação em Ciência da Computação) – Curso de Ciência da Computação, Universidade Federal do Paraná, Paraná, 2002.

KRISHNAMURTHY, Balachander. **Redes para Web.** Rio de Janeiro: Campus, 2001. 651 p.

KUROSE, James F.; ROSS, Keith W. **Redes de Computadores e a Internet:** uma abordagem top-down. 3. ed. São Paulo: Pearson Education do Brasil, 2006. 659p.

_____. **Redes de Computadores e a Internet:** uma abordagem top-down. 5. ed. São Paulo: Pearson Addison Wesley, 2010. 614 p.

LEVINE, David M. (Et al.). **Estatística - teoria e aplicações:** usando o microsoft Excel em português. Rio de Janeiro: LTC, 2012. 804 p.

LIMA, Janssen dos Reis. **Monitoramento de Redes com Zabbix:** monitore a saúde dos servidores e equipamentos de redes. Rio de Janeiro: Brasport, 2014.

LUCCA, Gustavo dos Santos de. **Diagnóstico do tráfego de rede web e análise da base de dados gerada pelo microsoft internet security and acceleration 2006:** estudo de caso na SATC. 2009. 110 f. Trabalho de Conclusão de Curso (Graduação) - Universidade do Extremo Sul Catarinense - Unesc, Criciúma, 2009.

MARTINS, Paulo João. **Comparação dos paradigmas cliente/servidor e agentes móveis:** um estudo em gerência de redes. 2002. 90 f. Dissertação (Mestrado) - Universidade Federal de Santa Catarina, Florianópolis, 2002.

MENESES, Anderson; MARIANO, Fabricio. **Noções de estatística para concursos.** Rio de Janeiro: Elsevier, 2010. 160p.

MILONE, Giuseppe. **Estatística:** geral e aplicada. São Paulo: Thomson, 2004. 483 p.

MONTGOMERY, Douglas C.; RUNGER, George C.; HUBELE, Norma Faris. **Estatística aplicada à engenharia.** 2.ed Rio de Janeiro: LTC, 2004. 335p.

MOTA FILHO, João E. M. **Pequenas redes com microsoft windows, para casa e escritório.** Rio de Janeiro: Ciência Moderna, 2001. 294p.

NASCIMENTO, Marcelo Brenzink do; TAVARES, Alexei Corrêa. **Roteadores e Switches:** guia de configuração para certificações CCNA. Rio de Janeiro: Ciência Moderna Ltda., 2006.

NORTHCUTT, Stephen (Et al.). **Desvendando:** segurança em redes. Rio de Janeiro: Campus, 2002. 650 p.

OLIVEIRA, Wilson José de. **Hacker: invasão e proteção.** Florianópolis: Visual Books, 2000. 108 p.

PETERSON, Larry L. **Redes de computadores:** uma abordagem de sistemas. Rio de Janeiro: Elsevier, 2004. 588 p.

POMPERMAYER JR, Jorge Luiz. **Protótipo de software para a monitoração de pacotes em uma rede tcp/ip em ambientes linux.** Blumenau: Universidade Regional de Blumenau - Centro de Ciências Exatas e Naturais - Curso de Ciências da Computação, 2002.

ROSS, Julio. **Redes de Computadores.** Cidade: Antenna Edições Técnicas, 2008. 148 p.

SOARES, Luiz Fernando Gomes; LEMOS, Guido; COLCHER, Sérgio. **Redes de computadores das LANs, MANs e WANs às Redes ATM.** 2. ed. Rio de Janeiro: Campus, 1995. 704 p.

SOUZA, Lindeberg Barros de. **Projetos e implementação de redes:** fundamentos, soluções, arquiteturas e planejamento. São Paulo: Érica, 2007. 320 p.

SPIEGEL, Murray R.; STEPHENS, Larry J. **Estatística.** 4 ed. Porto Alegre: Bookman, 2009. 597 p.

STALLINGS, Willinam. **Redes e sistemas de comunicação de dados:** teoria e aplicações corporativas. Rio de Janeiro: Elsevier, 2005. 449p.

TANENBAUM, Andrew S. **Redes de computadores.** 4.ed Rio de Janeiro: Campus, 2003. 632 p.

_____. **Redes de computadores.** 5.ed Rio de Janeiro: Campus, 2011. 923 p.

TAROUCO, Liane Margarida Rockenbach. **Redes de computadores locais e de longa distância.** São Paulo: Makron Books, 1986. 353 p.

TROMBIM, Diordgenes. **Diagnóstico do tráfego de rede de laboratórios de informatíca. Estudo de caso:** Universidade do Extremo Sul Catarinense. 2006. 116 f. Trabalho de Conclusão de Curso (Graduação) - Universidade do Extremo Sul Catarinense, Criciúma, 2006.

ULBRICH, Henrique César; VALLE, James Della. **Universidade Hacker:** desvende todos os segredos do submundo dos hackers. 5 ed. São Paulo: Editora Digerati, 2004. 348p.

APÊNDICE(S)

APÊNDICE A – ARTIGO

Análise de tráfego de rede

Eduardo S. Zardin[1]

[1]Curso de Ciência da Computação – Universidade do Extremo Sul Catarinense (UNESC) – Criciúma – SC – Brazil

eduardozardin@gmail.com

Abstract. *This article aims to analyze the traffic to existing networks at Universidade do Extremo Sul Catarinense. Data were obtained from a sniffer, in which each sample on statistical methods were applied to compare the results obtained in the periods analyzed. The results show that the bigger traffic generated by the administrative network was internally, concentrated in the morning and in the laboratories was for access to the Internet with greater flow at night. The obtained analyzes show that the purpose is to assist in the decisions that may be taken to obtain a network with better performance and flow of information.*

Resumo. *Este artigo tem o objetivo de analisar o tráfego nas redes existentes na Universidade do Extremo Sul Catarinense. Os dados foram obtidos a partir de um sniffer, no qual sobre cada amostra foram aplicados métodos estatísticos comparando os resultados obtidos nos períodos analisados. Os resultados demonstram que o maior tráfego gerado pela rede administrativa foi interno, concentrado no período matutino e na de laboratórios foi por acessos à Internet, com maior fluxo no período noturno. As análises obtidas demonstram que o propósito é auxiliar nas decisões que possam ser tomadas para obter-se uma rede com melhor desempenho e vazão de informação.*

1. Introdução

A infraestrutura de redes de computadores está ficando cada vez mais importante para as organizações e trata-se de um serviço indispensável que precisa estar 100% (cem por cento) operacional (LIMA, 2014).

Recursos computacionais nos permitem iniciar um trabalho no escritório, visualizá-lo no smartphone e finalizá-lo em casa no notebook ou no desktop. Com o avanço da tecnologia, é notável a necessidade de transmissão de informações entre diversos dispositivos tais como: desktops, notebooks, tablets e smartphones (COMER, 2006).

Com isso, Kurose e Ross (2006) destacam a necessidade do monitoramento e controle do fluxo de dados pois profissionais acabam desconhecendo o que trafega em sua própria rede. Como consequência, tomadas de decisões a respeito do desempenho de uma rede são definidas de forma errônea.

Levando em consideração essas informações, observa-se a necessidade de profissionais da área da tecnologia obterem um melhor detalhamento do tráfego de dados da rede. Analisar congestionamento de dados, identificando o causador de determinado problema que pode ser um usuário ou algo que está trafegando pela rede naquele momento, assim como os períodos mais críticos de acessos e a causa desses problemas.

Existem ferramentas de monitoramento que auxiliam esses profissionais, gerando informações dos dados obtidos, através de gráficos e outras formas de visualização, e um exemplo destas ferramentas é o sniffer que faz a captura de pacotes na rede e possibilita a verificação de seu conteúdo.

Nesta pesquisa foi feito uma avaliação na rede da Universidade do Extremo Sul Catarinense (UNESC) por meio de ferramentas de monitoramento de rede, com vistas a propor melhorias na infraestrutura.

2. Materiais e Métodos

Segundo Crespo (2009) quando necessita-se de uma conclusão sobre um todo (população), é normal observar-se esse todo na forma de amostras, ou seja, dividindo-o em partes.

Alguns trabalhos já realizados como o de Trombim (2006) e Jesus (2008) utilizando a técnica de amostragem estratificada, demonstraram sucesso em sua realização. Com base nesses dados, esta pesquisa propõe-se a utilizá-la.

Para complementar esta técnica, aplica-se também a amostragem casual ou aleatória simples, que consiste em numerar uma população de 1 a n em seguida sortear k números desta sequência, resultando nos elementos pertencentes à amostra (CRESPO, 2009).

Os métodos utilizados viabilizam a análise no tráfego de rede na Universidade do Extremo Sul Catarinense, pelo fato da pesquisa ser realizada em dois tipos de redes complexas e de grande porte com imenso volume de informações trafegando pela rede.

Foi desenvolvido uma metodologia para as capturas de pacotes, utilizando a técnica de amostragem estratificada para obter os períodos diários, em seguida aplica-se a técnica de amostragem aleatória para definição dos horários para coleta dos dados. Para tal, as horas foram convertidas em minutos e numeradas de 0 a 240, totalizando 4 horas em cada período definido como matutino, vespertino e noturno. O comando de sorteio utilizado com o auxílio do softwares Microsoft Office Excel versão 2013 foi “=aleatórioentre(0;240)”.

Para desenvolvimento deste artigo, utilizou-se a ferramenta Wireshark na versão 2.0.3 com a interface de rede do equipamento de coleta em modo promíscuo para realização da captura dos dados. Uma das principais características do Wireshark é sua interface gráfica, que facilita no momento da manipulação dos dados. Ressalta-se que essa ferramenta é open source.

Destaca-se ainda no software, o poder de identificação de vários pacotes que podem ser abertos visualmente, com a possibilidade de acompanhamento de todo o seu processo de encapsulamento.

3. Resultados

As análises estatísticas foram realizadas por meio do software IBM Statistical Package for the Social Sciencies (SPSS) versão 22.0. As variáveis foram expressas por meio de média, erro padrão, mediana e valores mínimos e máximos. Salienta-se também que alguns gráficos e tabelas foram criados utilizando-se recursos disponibilizados pela ferramenta Wireshark.

Conforme demonstra a Tabela 1, no intuito de se obter o tráfego total da rede administrativa e laboratório, aplicou-se um filtro na ferramenta Wireshark a fim de ignorar os protocolos ARP e RARP que são nativos da rede e que serão esboçados posteriormente.

Tabela 1 - Tráfego total da rede administrativa

Volume de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	300	24,20[a]	4,65	0,49	0	715,66	<0,001
Vespertino	300	20,49[a]	3,59	0,62	0	616,74	
Noturno	300	14,54[b]	2,60	0,20	0	405,81	

Fonte: do autor.
[a,b,c] –letras distintas representam diferenças estatisticamente significativas após a aplicação do teste Dunn (p<0,05).
* Valor obtido após aplicação do teste H de Kruskal-Wallis.

Na Tabela 1 observa-se que a rede administrativa apresentou uma variação em seu tráfego total, porém, manteve um comportamento padrão no período compreendido das 8h às 17h30, mantendo a média de tráfego entre 20 a 25 megabytes. No período noturno apresentou uma diferença estatisticamente significativa, com tráfego diminuído, chegando a atingir aproximadamente 10 megabytes a menos do que o decorrer do dia (p<0,001).

Outro ponto a ressaltar, foi quando o tráfego alcançou seu pico, atingindo cerca de 11,92 megabytes por segundo. Neste momento, aconteceu um comportamento no switch em que foi feita a captura diferente dos demais períodos de coleta, a retransmissão dos pacotes.

Analisando-se a Tabela 2, visualiza-se que o comportamento da rede laboratório foi o oposto da administrativa, demonstrando uma distinção estatisticamente significativa em seus três períodos durante a semana analisada, chegando a ter uma diferença de 10 megabytes entre eles e apresentando menor tráfego no período vespertino.

Vale ressaltar, que o maior pico de tráfego da rede laboratório ocorreu no período matutino, chegando a atingir cerca de 425 megabytes em apenas um minuto, mantendo a média de aproximadamente 10 megabytes. O maior volume aconteceu no período noturno com média em torno de 21 megabytes por segundo.

Tabela 2 - Tráfego total da rede laboratório

Volume de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	300	10,40[a]	2,49	0,15	0	424,64	<0,001
Vespertino	300	1,83[b]	0,43	0,09	0	80,65	
Noturno	300	20,61[c]	3,30	0,51	0	342,73	

Fonte: do autor.
[a,b,c] –letras distintas representam diferenças estatisticamente significativas após a aplicação do teste Dunn (p<0,05).

* Valor obtido após aplicação do teste H de Kruskal-Wallis.

A Tabela 3, indica que o tráfego de navegação WWW não apresentou uma diferença estatisticamente significativa (p<0,05). Nas médias apresentadas na tabela, estão o protocolo HTTP (transporte), e o TCP e UDP com destino a porta 80 (serviço World Wide Web).

Tabela 3 - Tráfego de Navegação (WWW) rede administrativa

Volume de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	20	4,36	0,95	3,35	1,03	20,75	<0,172
Vespertino	20	4,46	0,94	3,12	1,05	20,72	
Noturno	20	3,04	0,63	2,4	0,19	11,3	

Fonte: do autor.
* Valor obtido após aplicação do teste H de Kruskal-Wallis.

Conforme Tabela 4, os dados da navegação HTTPS mantiveram um comportamento padrão. Neste tráfego foram identificados os protocolos TCP com destino a porta 443, os SSL/TLS que são de segurança e que protegem a navegação por página e os serviços de e-mail.

Tabela 4 - Tráfego de navegação HTTPS rede administrativa

Volume de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	20	10,00	1,96	7,50	2,23	35,07	<0,088
Vespertino	20	8,25	1,28	7,05	2,72	30,22	
Noturno	20	6,12	0,96	5,24	0,82	19,88	

Fonte: do autor.
* Valor obtido após aplicação do teste H de Kruskal-Wallis.

Com a utilização de alguns recursos disponibilizados pela ferramenta Wireshark, observou-se que em horário de maior tráfego a navegação web do bloco S, representa em torno de 4% do tráfego total do ponto analisado e a HTTPS aproximadamente 4,9%.

Na Tabela 5 apresenta-se o tráfego interno gerado pelo bloco S. Identificou-se com maior incidência os protocolos de compartilhamento de arquivos SMB, de transporte TCP e UDP, o serviço de DNS, IP e o de gerência de rede SNMP, apontado como o terceiro com maior ocorrência na rede.

Constatou-se que no bloco S cerca de 86% do tráfego foi interno, identificando acesso aos servidores de arquivos, sistema acadêmico entre outros serviços.

Tabela 5 - Resumo do tráfego interno rede administrativa

Volume de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	20	12,16	3,26	7,46	2,46	57,98	<0,113
Vespertino	20	10,26	2,25	7,57	2,59	49,99	
Noturno	20	6,75	1,40	5,31	4,45	25,88	

Fonte: do autor.
* Valor obtido após aplicação do teste H de Kruskal-Wallis.

Prosseguindo-se com as análises, as tabelas 6 e 7, apresentam o tráfego de navegação web e HTTPS dos laboratórios e também outros resultados do bloco XXI-A.

Tabela 6 - Tráfego de navegação (WWW) rede laboratório

Volume de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	20	17,10[a]	5,74	7,13	0,76	94,42	<0,001
Vespertino	20	2,57[b]	1,01	0,88	0,03	20,36	
Noturno	20	29,48[a]	4,34	28,82	0,82	71,42	

Fonte: do autor.

[a,b] –letras distintas representam diferenças estatisticamente significativas após a aplicação do teste Dunn (p<0,05).

* Valor obtido após aplicação do teste H de Kruskal-Wallis.

Tabela 7 - Tráfego de navegação HTTPS rede laboratório

Volume de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*
Matutino	20	4,89[a]	1,36	2,40	0,60	24,54	<0,001
Vespertino	20	0,88[b]	0,23	0,52	0,15	0,47	
Noturno	20	10,24[c]	1,23	10,45	1,16	19,81	

Fonte: do autor.

[a,b,c] –letras distintas representam diferenças estatisticamente significativas após a aplicação do teste Dunn (p<0,05).

* Valor obtido após aplicação do teste H de Kruskal-Wallis.

Nas amostragens coletadas, identificou-se como principais protocolos o TCP direcionado para as portas 80 e 443; o UDP destinado a porta 80; o de transporte HTTP; os de segurança SSL/TLS e além dos já mencionados, o eXtensible Messaging and Presence Protocol (XMPP) definido pela Request for Comments (RFC) 3920 como protocolo aberto, que tem a finalidade de construir aplicativos de mensagens instantâneas. Este, não teve tanta incidência na rede administrativa, mas destacou-se na de laboratório.

Utilizando-se os recursos disponibilizados pela ferramenta Wireshark, identificou-se que na captura que apresentou maior tráfego a navegação web representou 28,3% e HTTPS representou 57,8% do tráfego, evidenciando que o maior tráfego da rede laboratório ocorre em acesso à Internet.

Realizou-se um filtro na ferramenta Wireshark para identificação do tráfego interno da rede laboratório, que é apresentado pela Tabela 8. Nesta análise aponta-se a presença dos serviços de DNS e DHCP, o protocolo TCP, o de gerência de rede SNMP e um outro que é integrante do IP definido pela RFC-792 (ICMP), que serve para fornecer relatórios de erros.

Tabela 8 - Tráfego interno rede laboratório

Volume de pacotes	N	Média	EP	Mediana	Mínimo	Máximo	Valor de P*

Matutino	20	0,62[a]	0,20	0,26	0,03	3,37	<0,001
Vespertino	20	0,10[b]	0,36	0,04	0	0,72	
Noturno	20	1,08[a]	0,16	1,06	0,04	2,56	

Fonte: do autor.

[a,b,c] –letras distintas representam diferenças estatisticamente significativas após a aplicação do teste Dunn ($p<0,05$).

* Valor obtido após aplicação do teste H de Kruskal-Wallis.

Observa-se na Tabela 8 que o comportamento do tráfego interno apresenta diferença entre os turnos, concentrando o maior número de pacotes capturados no período noturno. Em comparação com a rede do bloco S salienta-se que não teve tanta representatividade tendo gerado um tráfego de aproximadamente 1% do total, em que identificou-se acessos ao ambiente virtual, diário on-line e o monitoramento de ativos com o protocolo SNMP.

A partir dos dados obtidos com a análise da rede laboratório, evidencia-se que seu uso maior é em navegação web, concentrando a maior parte do tráfego no período noturno.

4. Discussão

No momento em que se desenvolve uma pesquisa num ambiente que está em atividade constante, são várias as dificuldades encontradas. Um dos primeiros problemas encontrados foi ao receber a informação de que não seria possível realizar a coleta dos dados no Armário de Telecomunicação principal da instituição, por políticas de segurança e por se tratar de um ambiente de produção. Nesse momento, surgiu a necessidade de estudar melhor toda a estrutura da rede e definir um novo ponto para fazer a coleta das informações.

Após a definição dos novos locais para análise, outra dificuldade surgiu, a forma de como deveria ser feito a coleta por ter um volume muito alto de dados. Com o auxílio de profissionais da área estatística, métodos foram definidos para que melhor se adequassem a situação e que validariam a pesquisa chegando o mais próximo do tráfego total possível.

O passo seguinte foi o estudo de uma alternativa para colocar a interface de rede do equipamento de coleta em modo promíscuo, para captura do tráfego geral nos pontos de coleta definidos, sendo que a rede da universidade é composta por Switchs, e com isso, disponibilizaria apenas o tráfego direcionado a este equipamento.

Depois de muito estudo e com auxílio dos profissionais da área de redes da instituição, foi encontrada a solução: como nos pontos de coletas os Switchs são gerenciáveis, foi realizado o direcionamento do tráfego da porta em que o link principal é ligado para alguma outra porta livre do próprio Switch, técnica conhecida como Port mirror. Contudo, como essa técnica nunca tinha sido utilizada, não havia conhecimento do que poderia ocorrer se aplicado com a rede em funcionamento.

Vários testes foram realizados em bancada, a fim de verificar problemas que pudessem ocorrer. Posteriormente, constatou-se que não teve alterações no comportamento do Switch, tanto no processamento como na memória. Então, foi definido um horário de menor pico na instituição para realização de um teste, que caso não resultasse positivamente, inviabilizaria a pesquisa.

Superadas as dificuldades apresentadas, foi realizada a coleta dos dados e por fim, a análise dos resultados obtidos, que possibilitaram o entendimento do comportamento das redes analisadas e o que trafega por elas.

5. Conclusão

Atualmente, a rede da UNESC comparada com outras empresas da região é consideravelmente de grande porte, tanto pela sua estrutura física, quanto ao parque de computadores existentes. Por meio de um software para monitoramento de ativos já utilizado pela instituição em seus servidores, observou-se que o comportamento de rede é heterogêneo durante o dia. Isto se deve ao fato da quantidade de usuários existentes e aos diversos perfis que utilizam o recurso. Foi a partir dessas informações que surgiu a ideia da realização desta pesquisa, que visou o estudo do tráfego e o comportamento das redes administrativa e de laboratórios existentes na instituição.

Destaca-se que o maior consumo da rede administrativa está relacionado aos acessos internos e na rede laboratório, evidencia-se que seu uso maior é em navegação web no período noturno. Além disso, identifica-se tráfego direcionado a mensageiros instantâneos tendo como principal mensageiro o Whatsapp e Snapchat. Visualizou-se alto índice de streaming de vídeo, com principal uso pelo Youtube e alguns sites de notícias, além de downloads de diversos tipos de mídia.

Existem vários estudos na comunidade científica abordando esse assunto e trabalhos já realizados que possuem semelhança com a presente pesquisa, contudo, aponta-se como diferença mais significativa em relação a essa, o volume de dados, a quantidade de equipamentos e também o maior número de amostragens obtidas, pois trata-se da análise de duas redes, a de laboratórios e uma administrativa. É válido ressaltar, que em relação as pesquisas já realizadas sobre o assunto abordado, foi acrescentada a análise de um outro ponto da rede da instituição, além da que já tinha sido estudada e a utilização de métodos estatísticos para apresentação dos dados, possibilitando um diagnóstico completo de toda a rede da Universidade. Esta pesquisa demonstra de forma resumida, aos profissionais da área, o comportamento das duas redes existentes.

Com os dados obtidos, destaca-se o comportamento distinto entre as duas redes, mas que possuem as mesmas características na sua arquitetura e funcionalidade.

Referências

COMER, Douglas. Interligação em rede com TCP/IP. Rio de janeiro: Elsevier, 2006.

CRESPO, Antônio Arnot. Estatística fácil. 19. ed. São Paulo: Saraiva, 2009. 218 p.

JESUS, Fabricio Cardoso de. Análise da rede sob o ponto de vista do controle de informações e tráfego estudo de caso: TSA Química do Brasil. 2008. 85 f. Trabalho de Conclusão de Curso (Graduação) - Universidade do Extremo Sul Catarinense - Unesc, Criciúma, 2008.

KUROSE, James F.; ROSS, Keith W. Redes de Computadores e a Internet: uma abordagem top-down. 3. ed. São Paulo: Pearson Education do Brasil, 2006. 659p.

LIMA, Janssen dos Reis. Monitoramento de Redes com Zabbix: monitore a saúde dos servidores e equipamentos de redes. Rio de Janeiro: Brasport, 2014.

TROMBIM, Diordgenes. Diagnóstico do tráfego de rede de laboratórios de informatíca. Estudo de caso: Universidade do Extremo Sul Catarinense. 2006. 116 f. Trabalho de Conclusão de Curso (Graduação) - Universidade do Extremo Sul Catarinense, Criciúma, 2006.

ANEXO(S)

ANEXO A – AUTORIZAÇÃO DE CAPTURA DE DADOS

unesc
Universidade
do Extremo
Sul Catarinense

AUTORIZAÇÃO DE CAPTURA DE DADOS

Declaramos para os devidos fins, que autorizamos o acadêmico **Eduardo de Souza Zardin** a realizar captura de dados (uso de *sniffer*), na rede Unesc, com o objetivo único de uso no seu Trabalho de Conclusão de Curso intitulado ANÁLISE DE TRÁFEGO DE DADOS EM REDES LOCAIS: ESTUDO DE CASO NA UNESC, sob a orientação do Prof. MSc. Rogério Antônio Casagrande.

Os pontos definidos para a coleta são o subsolo do bloco S e bloco XXI-A sala 9, garantindo a cobertura da rede administrativa e rede de laboratórios, respectivamente. O período previsto para realização da ação é de 02 a 06/05/2016.

Ressaltamos que os dados têm caráter confidencial e devem ser utilizados somente para os fins que se propõe a coleta, e ainda assim, serem divulgados de forma a não fragilizar a segurança do ambiente corporativo.

Criciúma, 7 de março de 2016

Valéria de Araújo
Gerente do Departamento de Tecnologia da Informação
Portaria nº 01/2016/Reitoria

Valéria de Araujo
Gerente do Departamento de Tecnologia da Informação