UNIVERSIDADE FEDERAL FLUMINENSE

FERNANDO EUGÊNIO MARCOS TEIXEIRA
FERNANDO SÉRGIO CARDOSO CUNHA
HENRIQUE SANTOS FERNANDES

ESTUDO DE VIABILIDADE DA IMPLEMENTAÇÃO DE UM PROTOCOLO DE PROTEÇÃO DE ANEL NO NÚCLEO DA REDE DA UNIVERSIDADE FEDERAL FLUMINENSE

NITERÓI

2014

UNIVERSIDADE FEDERAL FLUMINENSE

FERNANDO EUGÊNIO MARCOS TEIXEIRA
FERNANDO SÉRGIO CARDOSO CUNHA
HENRIQUE SANTOS FERNANDES

ESTUDO DE VIABILIDADE DA IMPLEMENTAÇÃO DE UM PROTOCOLO DE PROTEÇÃO DE ANEL NO NÚCLEO DA REDE DA UNIVERSIDADE FEDERAL FLUMINENSE

Trabalho de conclusão de curso orientado pelo professor João Marcos Meirelles da Silva e coorientado pela professora Natalia Castro Fernandes ambos integrantes do departamento de engenharia de telecomunicações da Universidade Federal Fluminense. Apresentado para banca examinadora no dia 15 de Janeiro de 2014

Orientador:
João Marcos Meirelles da Silva

Coorientadora:
Natalia Castro Fernandes

NITERÓI

2014

Fernando Eugênio Marcos Teixeira
Fernando Sérgio Cardoso Cunha
Henrique Santos Fernandes

ESTUDO DE VIABILIDADE DA IMPLEMENTAÇÃO DE UM PROTOCOLO DE PROTEÇÃO DE ANEL NO NÚCLEO DA REDE DA UNIVERSIDADE FEDERAL FLUMINENSE

Trabalho de conclusão de curso orientado pelo professor João Marcos Meirelles da Silva e coorientado pela professora Natalia Castro Fernandes ambos integrantes do departamento de engenharia de telecomunicações da Universidade Federal Fluminense. Apresentado para banca examinadora no dia 15 de Janeiro de 2014

Aprovada em JANEIRO de 2014.

BANCA EXAMINADORA

Prof. Dr. João Marcos Meirelles da Silva - Orientador

Prof. Dra. Natalia Castro Fernandes - Coorientadora

Prof. Dr. Ricardo Campanha Carrano

Prof. Dr. Tadeu Nagashima Ferreira

Niterói
2014

Dedico este trabalho a minha familía que me forneceu os meios necessários para lutar por meus objetivos e a meus amigos, que encheram esta jornada com experiências únicas.

Fernando Eugênio Marcos Teixeira

Primeiramente a Deus por me guiar ao longo deste caminho. A minha família por todo o apoio dado a mim durante todos os anos do curso. Aos meus amigos por todos os momentos memoráveis.

Fernando Sérgio Cardoso Cunha

Dedico este trabalho a minha família e meus amigos, principalmente aqueles amigos que todo ano me zoavam por eu nunca acabar a faculdade.

Henrique Santos Fernandes

Agradecimentos

Agradecemos aos professores João Marcos Meirelles da Silva e Natalia Castro Fernandes por toda orientação que nos deram nesse final de nossa formação como Engenheiros de Telecomunicações. Também agradecemos a Superintendência de Tecnologia e Informaçao da Universidade Federal Fluminense por prover todos os equipamentos e infraestrutura necessária para os experimentos do nosso trabalho.

Fernando Eugênio Marcos Teixeira, Fernando Sergio Cunha e Henrique Santos Fernandes.

Resumo

Hoje em dia, a quantidade de dados que trafegam pelas redes é imensa (e continua crescendo). Por esse motivo, as redes precisam de mecanismos para garantir alta disponibilidade de seus serviços. Nesse cenário, as redes com topologia anel ganham importância, pois a topologia em anel oferece redundância à um custo relativamente baixo. A Universidade Federal Fluminense possui uma rede de dados com topologia em anel de grande porte que oferece diversos tipos de serviços aos seus usuários. Dessa forma, o bom funcionamento do núcleo da rede de dados da UFF é de extrema importância para o bom andamento do ensino e pesquisa. Apesar da topologia em anel ser muito utilizada, ela gera *loops* na rede. Considerando esses fatores, esse trabalho estuda um protocolo de proteção de anel para ser implementado na rede de dados da Universidade Federal Fluminense e assim, evitar os *loops* e permitir a redundância. Outra peça importante para o bom desempenho de uma rede de dados é o monitoramento eficaz. Por isso, neste trabalho também é estudado e proposto o uso de ferramentas livres e de código aberto para que se possa implementar um monitoramento de qualidade.

Palavras-chave: UFF, STI, Anel, Resiliência, Redundância, RRPP, Monitoramento, Redes.

Abstract

Nowadays, the amount of data flowing through networks is huge (and growing). Therefore, the networks need mechanisms to ensure the high availability of it services. In this scenario, the networks with ring topology gain importance, because the ring topology offers redundancy at a relatively low cost. Universidade Federal Fluminense has a large data network with ring topology that offers various services to its users. Thus, the proper functioning of the core data of UFF's network is extremely important for the smooth progress of teaching and research. Although the ring topology is well used, it generates loops in the network. Considering these factors, this paper studies a protocol for ring protection to be implemented in the data network of Universidade Federal Fluminense and thus it is able to avoid loops and allow redundancy. Another key to the performance of a data network is the effective monitoring. Therefore, this paper also studies and proposes the use of free and open source tools to implement a quality monitoring.

Keywords: UFF, STI, Ring, Resiliency, Redundancy, RRPP, Networks, Network Monitoring.

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas e Siglas

STI	:	Superintendência de Tcnologia da Informação;
UFF	:	Iniversidade Federal Fluminense;
RRPP	:	Rapid Ring Protection Protocol;
IEEE	:	Institute of Electrical and Electronics Engineers;
VLAN	:	Virtual Local Area Network;
ERPS	:	Ethernet Ring Protection Switching;
REP	:	Resilient Ethernet Protocol;
RSTP	:	Rapid Spanning Tree Protocol;
MSTP	:	Multiple Spanning Tree Protocol;
SNMP	:	Simple Network Management Protocol;
SFP	:	Small Form-factor Pluggables;
QoE	:	Quality of Experience;
pop	:	points of presence;
SMS	:	Short Message Service;
ICMP	:	Internet Control Message Protocol;
DNS	:	Domain Name System;
LDAP	:	Lightweigth Directory Access Protocol;
NMS	:	Network Management System;
NTP	:	Network Time Protocol;
ISP	:	Internet Service Provider;
DHCP	:	Dynamic Host Configuration Protocol;
QoS	:	Quality Of Service;

Sumário

Capítulo 1

Introdução

A terminologia "rede" pode ser utilizada para qualquer tipo de conjunto de ligações físicas que se destinam a um mesmo serviço. Estes sistemas podem ser de diversos tipos: sistemas de satélites, sistemas de rádio visibilidade, redes de telefonia e redes de dados.

A implementação de uma rede eficaz tornou-se crucial para suportar a demanda que as aplicações estão gerando. Além disso, a convergência entre as redes chegou a tal ponto que hoje em dia não se consegue segregar uma rede telefônica de uma rede de dados. A antiga telefonia baseada em comutação por circuitos está gradativamente sendo substituída por uma nova geração que trafega quase que integralmente por redes de comutação de pacotes.

Com esse e muitos outros pontos de atuação, as redes de telecomunicações se tornaram uma das instalações de infraestrutura básica para a sociedade atual e sua dinâmica. Com o objetivo de aplicar grande parte dos conhecimentos voltados a essa área que o curso de Engenharia de Telecomunicações proporciona, este trabalho expõe uma realidade e as atitudes que foram tomadas em conjunto com a equipe da Superintendência de Tecnologia e Informação (STI) da Universidade Federal Fluminense (UFF) para adequar a estrutura que contempla os alunos e docentes em todos os campi da universidade, principalmente os localizados em Níterói, podendo assim prover um serviço de melhor qualidade.

O presente capítulo trata dos fatores motivacionais, da importância e do objetivo da elaboração deste trabalho.

O Capítulo 2 apresenta uma visão geral a respeito de RRPP. Apresenta seu funcionamento, dá detalhes a respeito da nomenclatura utilizada, quais são as aplicações do mesmo e apresenta sua configuração escolhida para este trabalho.

O Capítulo 3 explica a importância do monitoramento de redes assim como apresenta

ferramentas utilizadas e suas principais áreas de atuação.

O Capítulo 4 expõe a atual situação da rede de dados da UFF, explicitando sua topologia e exemplificando alguns de seus problemas.

O Capítulo 5 faz a proposta para o núcleo da rede de dados da UFF para solucionar alguns problemas conhecidos.

O Capitulo 6 descreve toda a metodologia utilizada para realização dos experimentos e traz também parte dos resultados encontrados.

1.1 Motivação

A demanda cada vez maior de banda de comunicação por parte de usuários, bem como o volume de negócios e serviços que hoje dependem da infraestrutura de rede, exerce uma pressão cada vez mais intensa sobre a disponibilidade e a confiabilidade destas. Novos protocolos tornam-se necessários para não só o transporte de dados mas para auxiliar na supervisão e gerenciamento de redes, permitindo assim uma maior redundância e consequentemente menos interrupções.

Nesse contexto, existe na UFF a necessidade de reestruturação do núcleo da rede de dados e a adequação a algumas exigências para a utilização da mesma para o trafego de voz, de forma a melhor atender os seus usuários.

Uma proposta técnica foi apresentada a STI, com a intenção de contestar a sua viabilidade técnica, experimentos foram realizados e seus resultados estão expostos ao longo deste trabalho.

1.2 Objetivo

Este trabalho tem como objetivo principal analisar a configuração atual do núcleo da rede de dados da UFF, testar o protocolo de proteção de anel *Rapid Ring Protection Protocol* (RRPP) e propor uma nova solução apoiada nos testes para a rede de dados. Além disso, esse trabalho visa mostrar que, devido às redundâncias, o monitoramento se torna critico para o bom funcionamento da rede.

Mostra-se também que o RRPP protege o anel de uma falha em um único enlace, com um tempo de convergência diminuto fazendo com que esta falha não afete a experiência do usuário.

Capítulo 2

Rapid Ring Protection Protocol - RRPP

O RRPP é um protocolo L2 de proteção para redes em anel e é proprietário da H3C. A seguir o seu funcionamento é descrito de forma que o resto deste trabalho possa ser melhor compreendido [1].

2.1 Visão Geral

2.1.1 Introdução

A topologia em anel tem sido amplamente utilizada por causa de sua confiabilidade. Porém, um problema recorrente e muito prejudicial a esse tipo de rede é o *loops*. Os protocolos desenvolvidos pelo *Institute of Electrical and Electronics Engineers* (IEEE) são muito utilizados para prevenção de *loops* [2]. Mas conforme a rede cresce, o tempo de convergência também cresce. Para sanar este problema, a H3C desenvolveu o RRPP.

2.2 Conceitos Básicos

2.2.1 Porta Primária e Porta Secundária

A escolha de quais interfaces serão porta primária e secundária ocorre quando o *Master node* envia pela primeira vez a mensagem *HELLO*. A porta primária de cada *transit node* é aquela que recebe a mensagem *HELLO* e a secundária é aquela passa a mesma mensagem adiante. Para o caso do *Master node*, a porta primária é aquela que gera a mensagem *HELLO* e a secundária é a que "espera" a mensagem *HELLO* rodar o anel inteiro.

Os conceitos de *Master node*, *transit node* e a mensagem *HELLO* são explicados adiante neste mesmo Capítulo 2.

2.2.2 Porta RRPP e Portas Não RRPP

As portas RRPP dos *switches* são aquelas portas que participam de um anel RRPP. As portas não RRPP por sua vez, são aquelas que não fazem parte de um anel RRPP.

2.2.3 Domínio

Um domínio RRPP define a topologia que será controlada pelo protocolo e é identificado por um número inteiro chamado de ID. É composto por dispositivos interconectados com um mesmo "*domain* ID", a mesma *Virtual Local Area Network* (VLAN) de proteção e mesmas VLANs de controle (mudam apenas entre os anéis que compõe o domínio). Um dispositivo pode pertencer a diversos domínios RRPP.

Os elementos que compõe um mesmo domínio são:

- Anéis RRPP
- VLANs de controle
- VLANs de proteção
- *Master nodes*
- *Transit nodes*

Na Figura 2.1, tem-se um domínio RRPP chamado de "*Domain* 1". Ele contém o anel 1 e dispositivos de S1 a S4. O anel 1 possui um *master node* (S1) e três *transit nodes* (S2, S3 e S4). A VLAN de controle primária é a 3. Esses elementos são explicados adiante.

2.2.4 Anel Simples RRPP

Um anel simples RRPP é um anel Ethernet e é identificado por um ID[1]. Os cálculos do protocolo são feitos com base no anel. Para o anel simples, o RRPP deve ser configurado com cada elemento no mesmo domínio.

[1]Apesar de terem o mesmo nome, os IDs do anel RRPP e do domínio RRPP não são o mesmo.

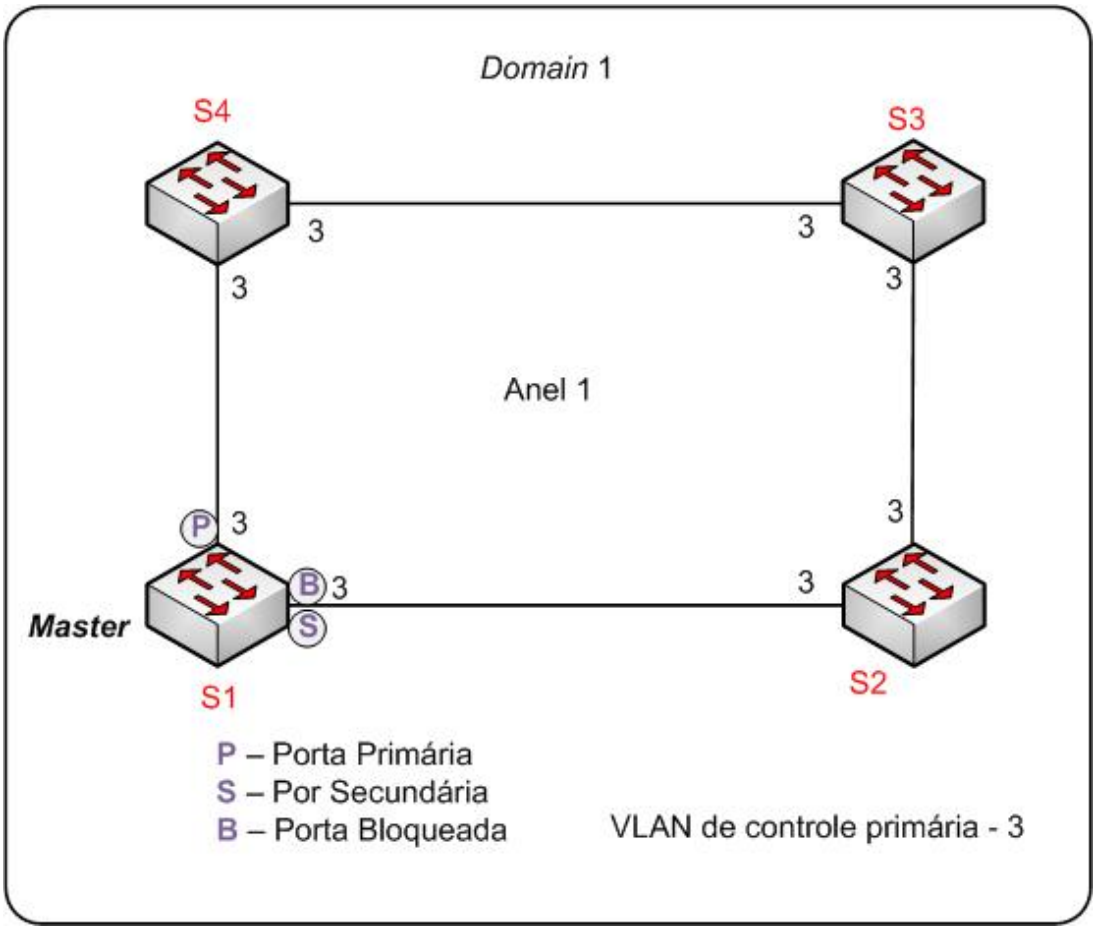

Figura 2.1: Domínio RRPP

2.2.5 VLAN de Proteção

A VLAN de proteção é a responsável por carregar os pacotes de dados. Ela não pode ser configurada com o mesmo número que a VLAN de controle, mas pode ser configurada em portas de RRPP e portas não RRPP. Para garantir o isolamento entre os domínios, cada um deverá ter um conjunto único de VLANs de proteção.

2.2.6 VLAN de Controle

As VLANs de controle são responsáveis pelo transporte das RRPPDUs. Todo domínio deve ser configurado com uma VLANs de controle.

Apenas dispositivos que estejam voltados para algum anel RRPP devem ter suas portas configuradas com VLANs de proteção. Na Figura 2.1 pode-se ver que os elementos do anel têm suas portas configuradas com a VLAN 3.

2.2.7 *Master Node*

O *master node* é o elemento responsável pelo controle do estado do anel. Cada anel deve ter um (e somente um) elemento configurado como *master node*.

Os estados possíveis para o *master* node são:

- *Complete State*

  O *master* node envia uma mensagem *HELLO* pela sua porta primária que deve rodar todo o anel. Caso esta mensagem chegue na sua porta secundária, o *master* node "sabe" que o anel está fechado e sem falhas. Assim, as VLANs de dados da porta secundária são bloqueadas para a prevenção de *loops.*

- *Failed State*

  Quando o *master* node envia a mensagem *HELLO* pela porta sua primária e esta não chega à porta secundária, o anel contém falha(s). Para que o fluxo de dados não seja interrompido o *master* node desbloqueia as VLANs de proteção da sua porta secundária.

2.2.8 *Transit Node*

O *transit node* é um elemento de apoio ao *master node*. Suas funções são simples: encaminhar as RRPPDUs, monitorar seus *links* RRPP diretamente conectados e alertar ao *master node* caso algum desses *links* mude de estado.

Um *transit node* também possui estados, que são:

- *Link-up State*

  Quando as portas primárias e secundárias do *transit node* estão transmitindo, então o mesmo encontra-se em *link-up state*.

- *Link-down State*

  O *transit node* encontra-se neste estado quando ou a porta primária ou a porta secundária estão em *link-down state*.

- *Pre-Forwarding State*

  Este estado ocorre quando o *transit node* detecta a recuperação de em um *link* conectado diretamente e passa a porta primária ou secundária do *transit node* para o estado de bloqueada.

2.2.9 Tipos de RRPPDUs

Os tipos de RRPPDus estão apresentados na Tabela 2.1.

Tabela 2.1: Tipos de RRPPDUs

HELLO	Mensagem enviada regularmente pelo *master node* para a verificação do estado do anel. Caso esta chegue na porta secundária do *master node*, o anel é considerado como fechado; se não, o anel é considerado em estado aberto.
LINK-DOWN	Usada pelo *transit node* para comunicar falhas de *links* ao *master node*.
COMMON-FLUSH-FDB	Esta mensagem é enviada pelo *master node* aos *transit nodes* quando o *master node* entra em estado de falha. Ao recebe-la, os *transit nodes* atualizam as entradas nas tabelas MAC e ARP/ND.
COMPLETE-FLUSH-FDB	O *master node* envia esta mensagem para notificar os *transit nodes* que o estado fechado foi alcançado. Com isso, os *transit nodes* atualizam suas tabelas MAC e ARP/ND, desbloqueiam as portas bloqueadas temporariamente e assumem o estado *link-up*.

2.3 Domínio único e Anel Simples

Existem três tipos básicos de topologias para o RRPP. Este trabalho fará uma abordagem mais detalhada apenas da topologia do anel simples, pois esta é a topologia proposta para o núcleo da rede IP da Universidade Federal Fluminense.

2.3.1 Funcionamento

O funcionamento do RRPP é divido em três etapas: anel completo, detecção de falha e recuperação de falha.

2.3.1.1 Anel Completo

O *master node* é o elemento chefe do anel RPPP. Ele envia periodicamente mensagens *HELLO* que devem rodar o anel para a verificação de falhas, determinação do estado do anel, prevenção de *loops* e de falha na transmissão de dados. A Figura 2.2 representa o anel

RRPP em *complete state*. Neste estado, o *master node* bloqueia a sua porta secundária e todos os *transit nodes* possuem as portas primárias e secundárias desbloqueadas.

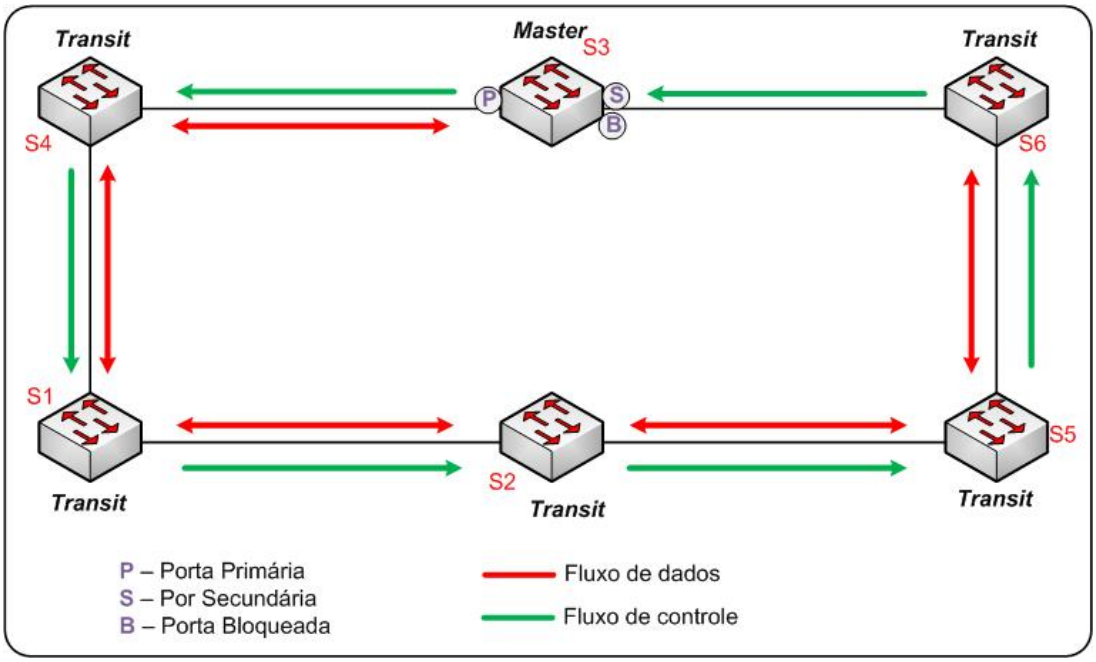

Figura 2.2: Anel RRPP em *complete state*

Como visto na Figura 2.2, o fluxo de controle segue apenas uma direção e o fluxo de dados não passa pela porta secundária do *master node*; evitando assim *loops*.

2.3.1.2 Detecção de Falha

A Figura 2.3 mostra uma das formas de detecção de falhas em anel RRPP que é a falha da mensagem *HELLO*. Esta falha ocorre quando a mensagem *HELLO* não chega na porta secundária do *master node* após um período pré-determinado

A partir deste ponto o anel é considerado aberto e o *master node* passa para o estado *failed state* e desbloqueia sua porta secundária. As mensagens *COMPLETE-FLUSH-FDB* são enviadas pelas portas primária e secundária do *master node* ordenando que os *transit nodes* atualizem as tabelas MAC e ARP/ND, como visto na Figura 2.4.

Assim, todos os elementos do anel estarão cientes que há uma falha.

Outra forma de detecção de falhas é o monitoramento dos *links* conectados diretamente. Quando o *master node* detectar uma falha em sua porta primária, ele imediatamente desbloqueará a porta secundária e enviará mensagens do tipo *COMPLETE-FLUSH-FDB* para que os *transit nodes* atualizem as tabelas MAC e ARP/ND. Os *transit nodes*s também podem ter a iniciativa na detecção de uma falha no anel RRPP. Para

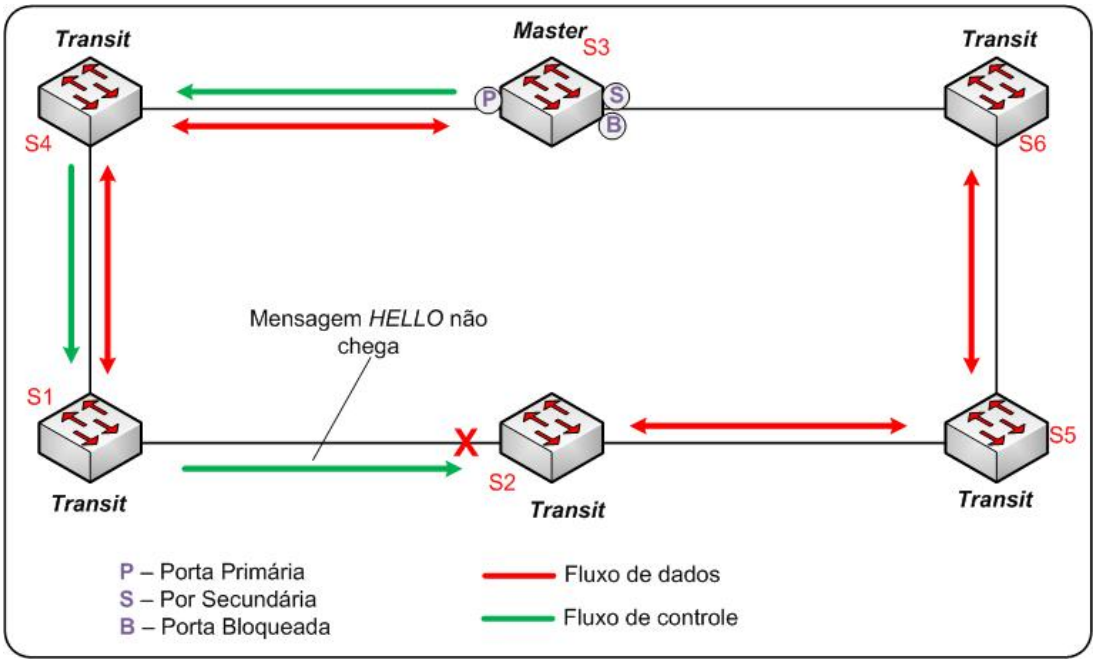

Figura 2.3: Falha da mensagem *HELLO* em um anel RRPP

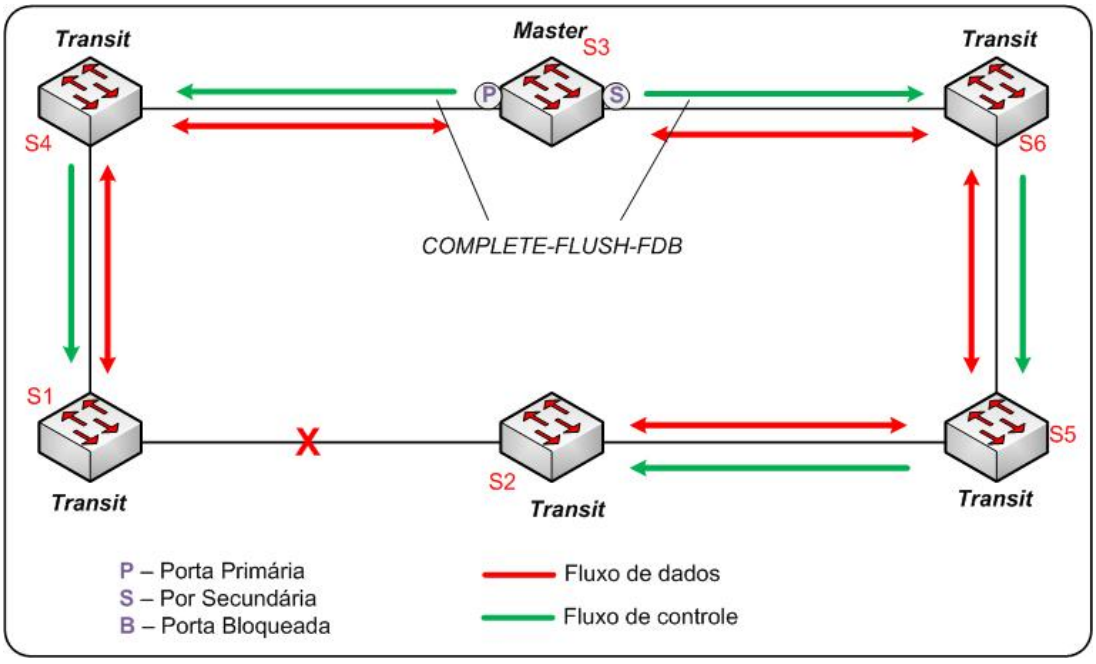

Figura 2.4: Mensagem RRPP *Complete Flush*

isso, o *transit node* deve detectar uma falha em algum *link* conectado diretamente e avisar ao *master node* através de uma mensagem do tipo *LINK-DOWN*, como mostrado na Figura 2.5.

Ao receber esta mensagem, o *master node* atualiza a sua própria tabela MAC, desbloqueia a sua porta secundária e envia aos *transit nodes* mensagens do tipo *COMMON-FLUSH-FDB* instruindo-os a atualizarem as suas tabelas MAC e ARP/ND, como mos-

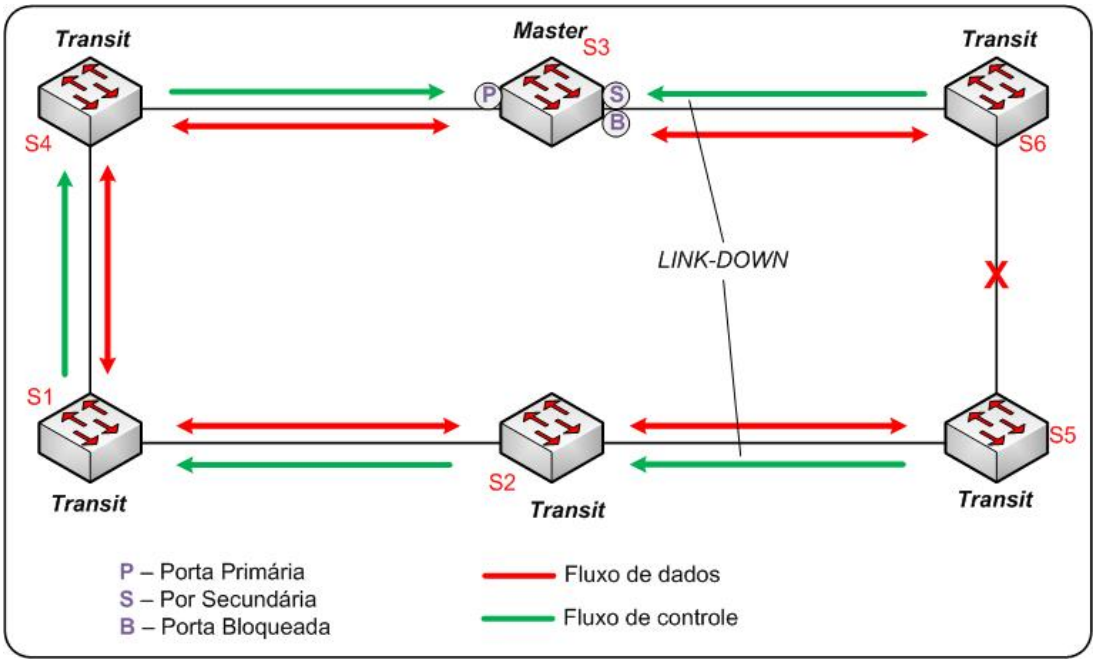

Figura 2.5: Mensagem RRPP *LINK-DOWN*

trado na Figura 2.6.

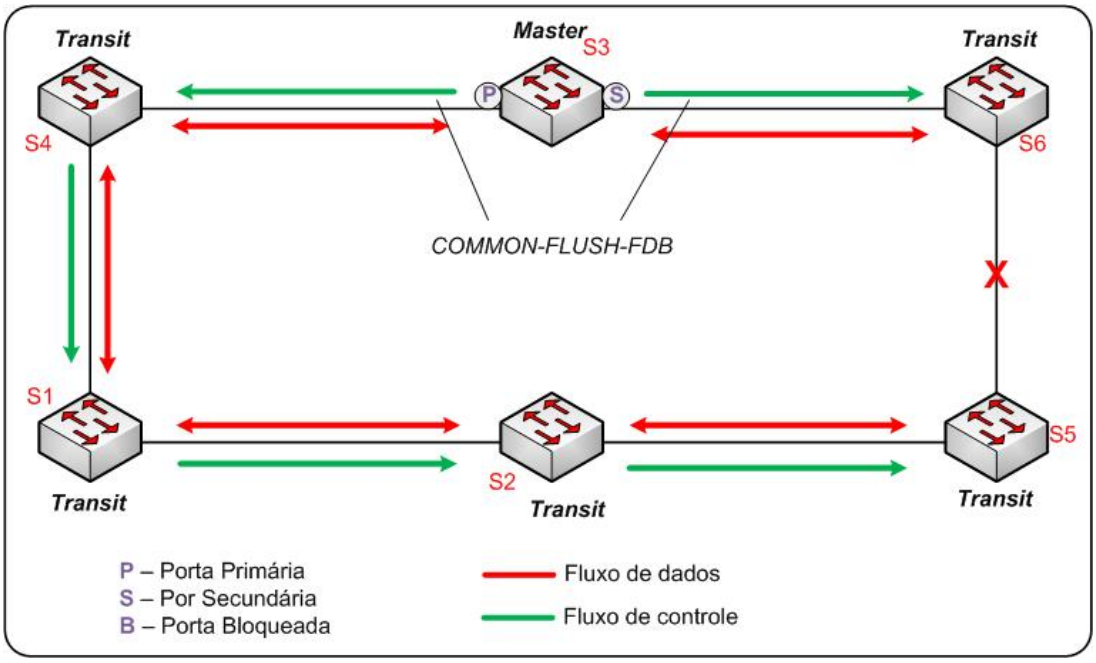

Figura 2.6: Mensagem RRPP *COMMON-FLUSH-FDB*

2.3.1.3 Recuperação de Falha

Quando um *link* é recuperado, o *master node* não é avisado imediatamente e consequentemente as suas portas primária e secundária estão desbloqueadas. Por isso, caso

os *transit nodes*, conectados diretamente ao *link* que estava com falha, abram as portas onde o *link* foi reestabelecido, ocorrerá *loops* até que o *master node* seja avisado. Para evitar esse *loops* momentâneo, os *transit nodes* (S5 e S6), ao detectarem a recuperação de um *link* conectado diretamente, passam para o estado *pre-forwarding* e bloqueiam a suas respectivas portas primária e secundária para que o *master node* possa inciar o processo de recuperação do anel, como visto na Figura 2.7.

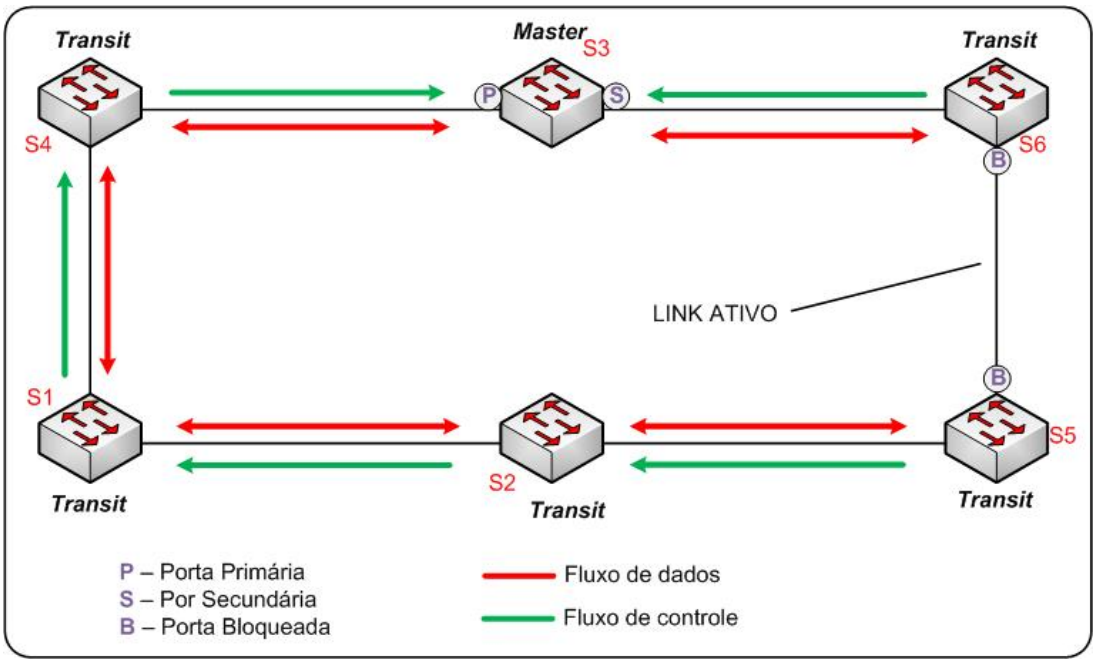

Figura 2.7: Detecção da recuperação de um *link* por um *transit node*.

Com o anel recuperado, o *master node* recebe novamente as mensagens *HELLO* em sua porta secundária. Ele, então, atualiza a sua tabela MAC e ARP/ND e envia mensagens do tipo *COMPLETE-FLUSH-FDB* para que todos os outros elementos do anel atualizem suas tabelas MAC e ARP/ND.

2.4 Outros Protocolos para Proteção de Redes em Anel

Existem diversos protocolos de proteção de redes em anel. Todos funcionam de forma semelhante: um dos elementos é escolhido para ser o mestre do anel e tomar as decisões para a prevenção de *loops* e perda de comunicação.

Alguns dos principais protocolos usados no mercado hoje em dia são:

- Ethernet Ring Protection Switching (ERPS) - recomendação ITU-T G.8032.[3]
- Resilient Ethernet Protocol (REP) - protocolo proprietário da Cisco.[4]

- Rapid Spanning Tree Protocol (RSTP) - IEEE 802.1w[5]
- Multiple Spanning Tree Protocol (MSTP) - IEEE 802.1s[6]

O Protocolo RRPP foi o escolhido, pois os equipamentos disponíveis para o projeto não são compatíveis com o ERPS e o tempo de convergência do RRPP (50 ms) é melhor do que o tempo de convergência dos protocolos RSTP e MSTP (no mínimo 1 s).

2.5 Outras Topologias RRPP

A topologia do anel simples não é a única com que o RRPP pode trabalhar. As outras duas topologias são: secante e tangente.

Para cada um desses tipos, o RRPP se comporta de maneira diferente:

- Para os anéis do tipo secantes, o RRPP deve ser configurado com cada elemento no mesmo domínio.
- Para o caso de anéis tangentes, cada anel deve ser configurado em um domínio diferente, ou seja, são necessários "n" IDs para o caso de "n" anéis tangentes.

O caso de anéis secantes (visto na Figura 2.8) merece uma atenção especial, pois, para que uma topologia independente seja alcançada, são necessárias configurações diferentes. Para se entender a topologia de anéis secantes, são necessários os conceitos de VLAN de controle primária, VLAN de controle secundária, anel primário e anel secundário. A VLAN de controle primária é a mesma utilizada na topologia de anel simples. Quando VLAN de controle primária é configurada, o *switch* configura automaticamente a VLAN secundária aumentando em um o número da VLAN primária. O anel primário é escolhido manualmente e deve ser configurado com nível 0 e nele as VLANs de controle primária e secundária estão configuradas nas portas RRPP. Os anéis restantes serão os secundários e deverão ser configurados com nível e neles, apenas a VLAN secundária está configurada nas portas RRPP.

As RRPPDUs (pacotes do protocolo RRPP) do anéis secundários são transportados de modo transparente no anel primário. Já as RRPPDUs do anel primário ficam restritas ao mesmo. Dessa forma, os anéis secundários enxergam o anel primário como mais um nó lógico da topologia; e assim o cálculo pode ser feito considerando a topologia como um todo.

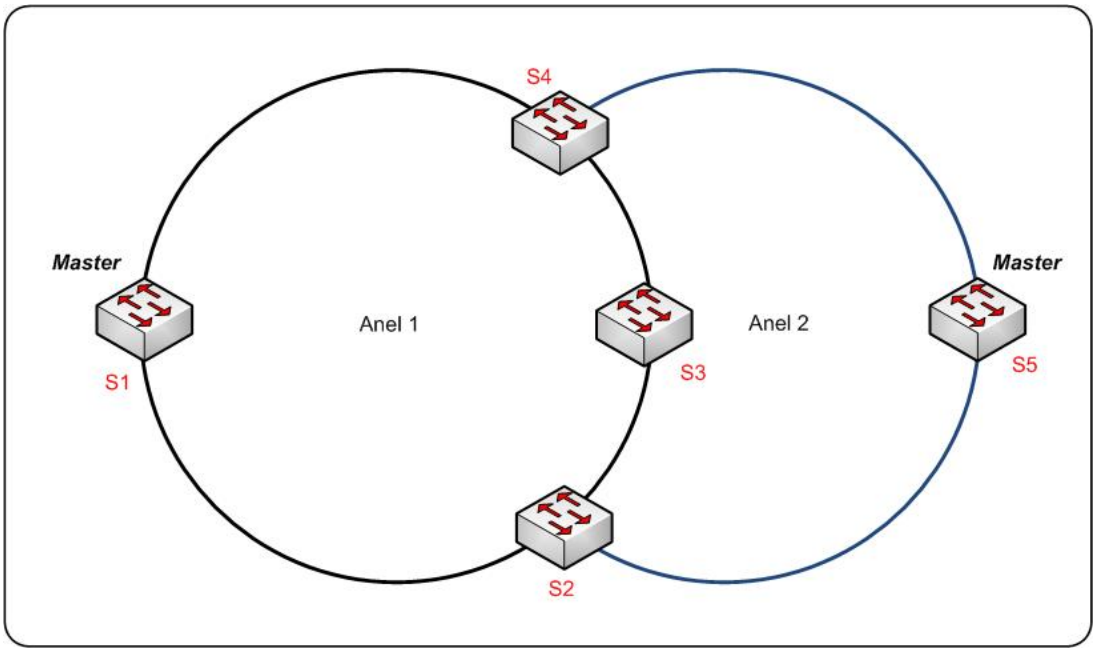

Figura 2.8: Topologia RRPP de anéis secantes.

O anel primário deve ser configurado com as VLANs primária e secundária e dessa forma, transportará as respectivas RRPPDUS e as mensagens de *EDGE- HELLO* do anel secundário.

Os anéis secundários devem ser configurados apenas com a VLAN secundária. Dessa forma, apenas as RRPPDUs dos subanéis serão transportadas, com exceção das mensagens *EDGE-HELLO*[2].

A topologia de anéis tangentes (vista na Figura 2.9) funciona de forma semelhante à de anel simples. A diferença é que o elemento que pertence aos dois domínios (*switch* S3 da Figura 2.9), tem quatro portas RRPP (duas para cada anel). Cada par de portas pertencentes ao mesmo anel, ou seja, mesmo domínio, deve ser configurada com a mesma VLAN de controle. As VLANs de controle de cada par devem ser diferentes entre si.

2.6 Benefícios

O RRPP é um protocolo de prevenção de *loops* dedicado à tecnologia Ethernet. Suas vantagens são:

- Convergência em até 50ms.
- Tempo de convergência independente do tamanho da rede.

[2]Mensagem que verifica o estado dos subanéis.

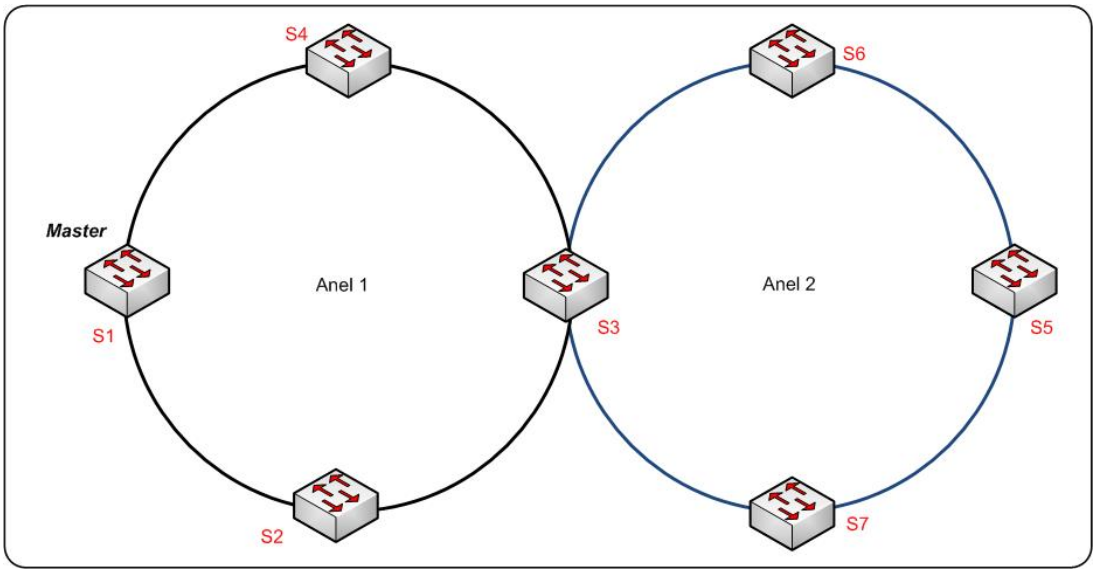

Figura 2.9: Topologia RRPP de anéis tangentes.

Em redes com múltiplos anéis a mudança na topologia de um anel com o RRPP não altera a topologia dos restantes, aumentando a estabilidade das conexões. O RRPP também suporta balanceamento de carga entre os anéis, otimizando assim a utilização de banda.

Capítulo 3

Tecnologias de Monitoramento da Rede

3.1 Introdução

Este capitulo tem como objetivo apresentar a importância do monitoramento efetivo de uma rede de computadores para a gerencia eficiente da mesma, evidenciando pontos de falha do cenário atual da rede que poderiam ser evitados caso fossem conhecidos. Além disso, tem-se uma proposta de implementação de uma estrutura de monitoramento que visa identificar tais falhas e propiciar um melhor desempenho.

Quando se trata de uma rede tão grande e heterogênea como a que contempla os campi da Universidade Federal Fluminense, composta por diferentes tipos de ativos (Roteadores e *Switches*) e servidores, de diferentes fabricantes, o monitoramento proprietário de um ou outro fornecedor se torna inviável. Sendo assim, para o recolhimento das informações, protocolos padronizados, que já tem sua aceitação no mercado e são comuns a todas as marcas, são valorizados. Para essa solução, em específico, utiliza-se o *Simple Network Management Protocol* (SNMP). Por se tratar de um protocolo comum a todos os fabricantes, a interoperabilidade das aplicações se dá com suavidade e poucos ajustes são necessários para casos específicos. Um tratamento singular das *Management Information Bases* (MIBs) se faz necessário, pois muitas vezes, elas diferem no que condiz ao formato escolhido por fabricante.

No cenário estudado, a existência de mais de quinhentos ativos compondo a rede e cerca de cem servidores com pouco ou quase nenhum monitoramento é um problema conhecido. A ineficiência da identificação de falhas e a impossibilidade de consultar um histórico de atividades da rede e dos próprios servidores gera um entrave para o diagnóstico e solução de qualquer problema.

Este caso deixa claro que o monitoramento e o acompanhamento das atividades viabilizaria a operação e resultaria em um número de falhas muito menor. Pontos de possível falha podem ser conhecidos e medidas preventivas podem ser tomadas visando a melhoria do desempenho de enlaces sobrecarregados, os quais causam uma *Quality of Experience* (QoE) indesejável para o usuário final.

O monitoramento pró-ativo é uma das formas de manter a confiança do usuário na rede intacta. Mesmo que algum evento imprevisível venha a acontecer, as equipes de suporte podem atuar assim que o mesmo ocorrer, se o sistema de alarmes estiver bem ajustado.

Por esses motivos, uma estrutura de monitoramento neste cenário se torna indispensável. A proposta desse trabalho se baseia integralmente em aplicações livres, desenvolvidas pela comunidade internacional, que contribui com novas atualizações e suporte. Apesar de serem ferramentas amplamente difundidas, nem todas possuem uma grande documentação online, e demandam um conhecimento básico em Linux. Abaixo, estão apresentadas as ferramentas utilizadas e sua área de atuação.

3.2 Ferramentas

3.2.1 Cacti

Cacti é uma aplicação com uma interface web desenvolvida para utilizar todo o poder oferecido pelas RRDTool's, ferramentas utilizadas para armazenamento de dados de forma mais eficiente para a geração de gráficos, como os vistos na Figura 3.1. Esses dados podem ser customizados pelo gestor de rede da forma que melhor atender as suas necessidades. Por se tratar de uma ferramenta livre, as contribuições de todos os usuários geram uma grande base de modelos de gráficos de diferentes tipos - tráfego, utilização de CPU, espaço em disco, uso de memória, *dhcp-pools* - que facilitam a utilização e potencializam esta aplicação.

A aquisição destas informações é feita de diversas formas que são igualmente maleáveis, mas o protocolo SNMP é mais comumente utilizado, por se tratar de um protocolo padronizado e não proprietário.

Além destas funcionalidades, a ferramenta também possui integração com extensões que podem adicionar outras funções, como a criação de mapas, alarmes de utilização excessiva de banda e alarmes de disponibilidade.

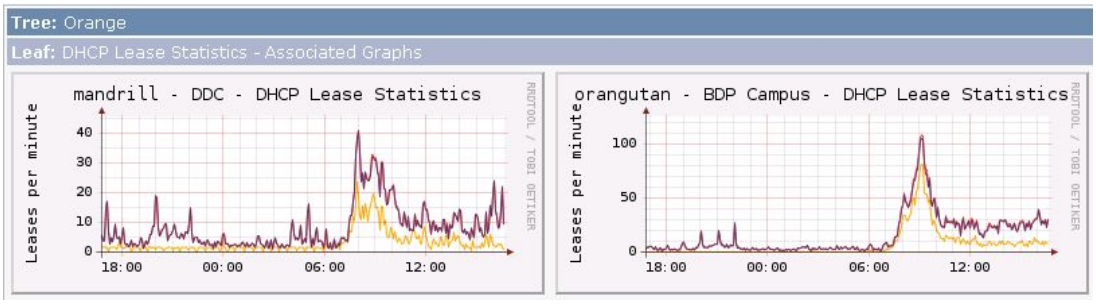

Figura 3.1: Exemplos de gráficos do Cacti

Há a possibilidade ainda de exportar os gráficos gerados para outros *websites*, podendo assim ter um portal de monitoramento único que possua a topologia e o panorama do uso atual da rede. A Rede Nacional de Ensino e Pesquisa (RNP) possui um portal deste tipo Figura 3.2 que demonstra a utilização dos *links* que ela administra entre seus *points of presence* (pop's).

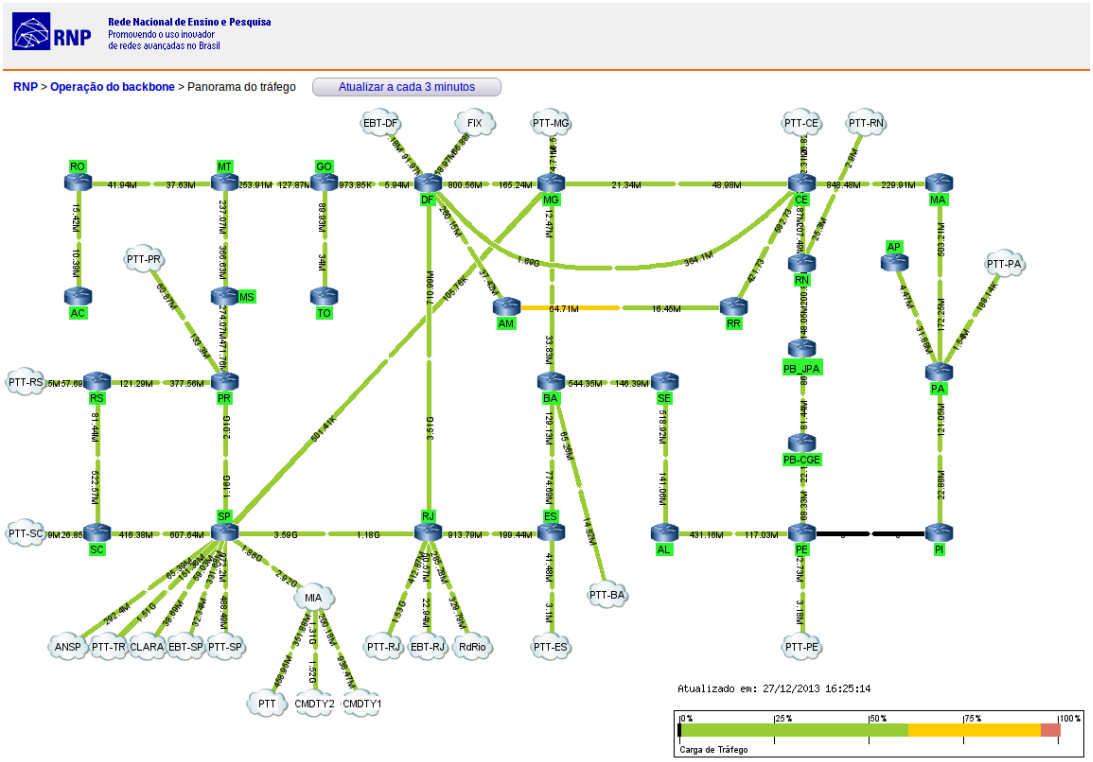

Figura 3.2: Exemplo do Weathermaps

3.2.2 Nagios

Nagios é uma ferramenta robusta para o monitoramento da disponibilidade de sistemas e aplicações, dando a possibilidade do gestor customizar os tipos de checagem que serão executados em cada *host* (servidor ou ativo), assim como qual alarme será dado e a

forma de aviso gerado ao responsável por tal serviço.

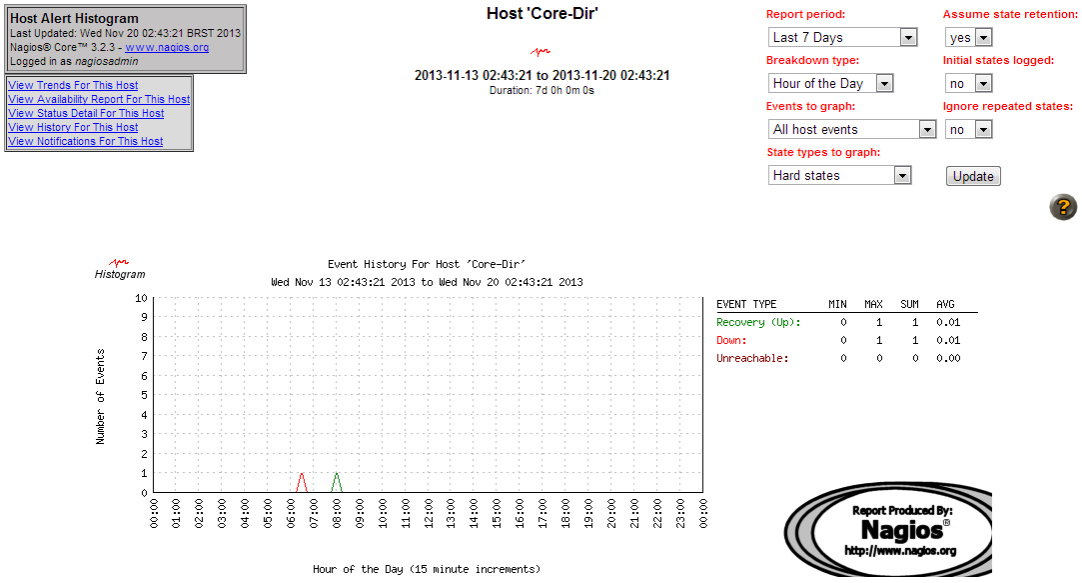

Figura 3.3: Exemplo de Histórico do Nagios

Assim como o cacti, ele possui um histórico de seus alarmes Figura 3.3 e uma interface *web* Figura 3.4 que facilita a visualização dos *hosts* e serviços atualmente monitorados.

As checagens executadas por esta ferramenta podem ser apenas um simples PING para confirmar a disponibilidade do *host* na rede ou um complexo *query* em uma tabela de uma base de dados, dependendo apenas do gestor, a forma e a necessidade de monitoramento do mesmo.

Assim como a maleabilidade das checagens, os alarmes podem ser desde simples envios de e-mail Figura 3.5 e *Short Message Service* (SMS), até a geração de uma ligação para um número de suporte para alertar o usuário de possíveis mensagens não visualizadas.

Ele proporciona também alguns parâmetros que podem gerar indicativos da qualidade dos enlaces no qual ele opera, pois avisos de alta latência e indisponibilidades momentâneas são igualmente importantes.

3.2.3 Smokeping

Quando uma rede passa a tomar grandes dimensões físicas, a manutenção da qualidade do serviço provido passa a ser tão importante quanto a disponibilidade do mesmo, pois, por muitas vezes, prover um serviço de baixa qualidade pode ser mais prejudicial do que não prover serviço algum.

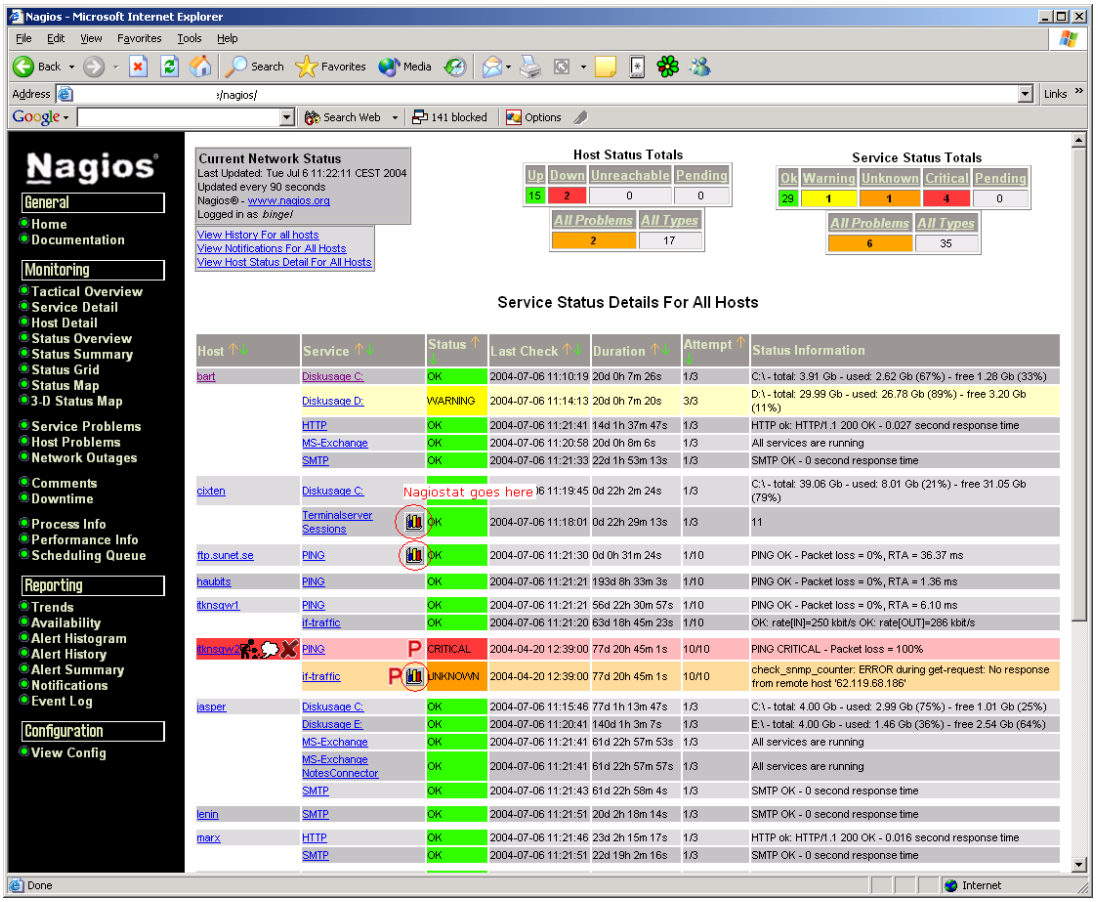

Figura 3.4: Exemplo da interface do Nagios

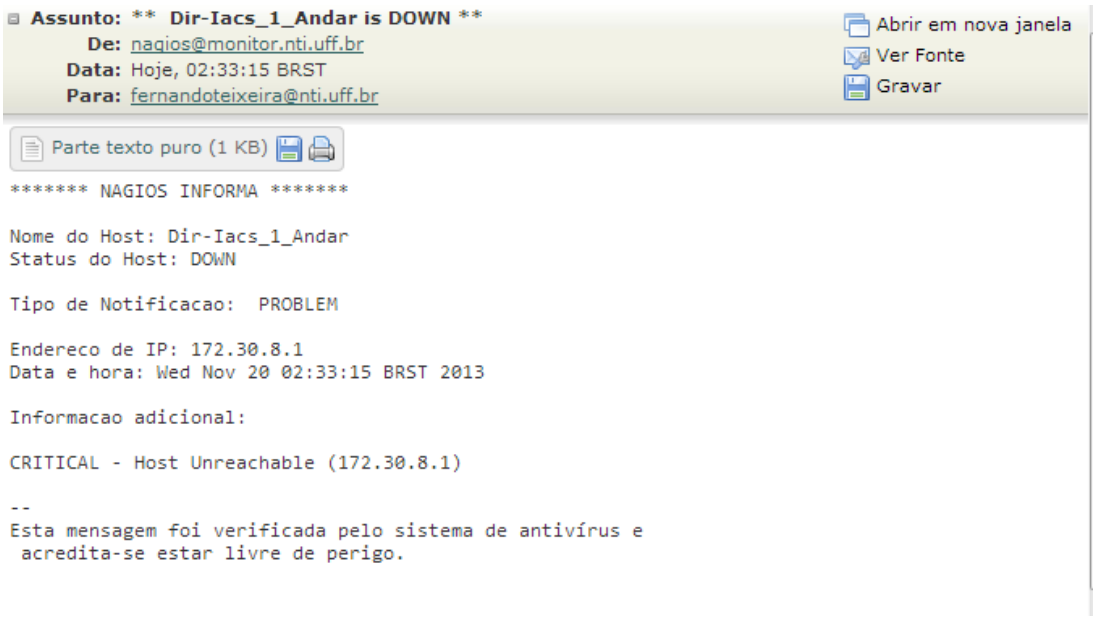

Figura 3.5: Exemplo de alerta do Nagios

O Smokeping foca na geração de históricos de latência entre o servidor, que normalmente é posicionado de forma estratégica na topologia da rede, próximo a serviços críticos

e a saída padrão da rede e os *hosts* especificados pelo gestor, podendo este ser qualquer tipo de ativo que tenha um IP.

Este servidor dispara pacotes *Internet Control Message Protocol* (ICMP) para os *hosts* selecionados em localidades diferentes na topologia descrita pelo gestor, e assim tem um panorama do desempenho que esta rede proporciona aos usuários durante todo o período de utilização e não apenas nos momentos que o usuário nota alguma instabilidade no serviço. Figura 3.6

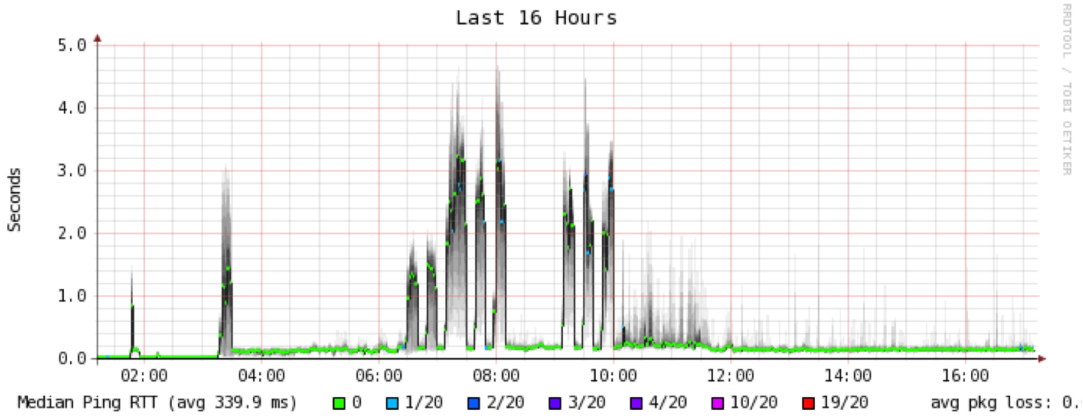

Figura 3.6: Exemplo de gráfico do Smokeping

Além das checagens de latência, ele possui facilidades e extensões prontas para capturar dados de outros tipos de serviço que são essenciais ao funcionamento da rede. Dentre eles temos: *Domain Name System* (DNS), *Lightweigth Directory Access Protocol* (LDAP) e tempo de resposta de aplicações *web*.

3.2.4 Rancid

O Rancid tem por objetivo principal a automatização do backup dos equipamentos, assim como o de aplicações. Ele possui uma interface web Figura 3.7 que facilita a visualização das configurações. Pode ser configurado para alarmar caso alguma mudança seja feita nas configurações de algum dos equipamentos cadastrados e registrar acessos indevidos.

Como funcionalidade extra, ele também oferece a possibilidade de configuração em escala dos equipamentos, facilitando assim a replicação de configurações padrão, como log remoto, *Network Management System* (NMS), *Network Time Protocol* (NTP) e formas de autenticação

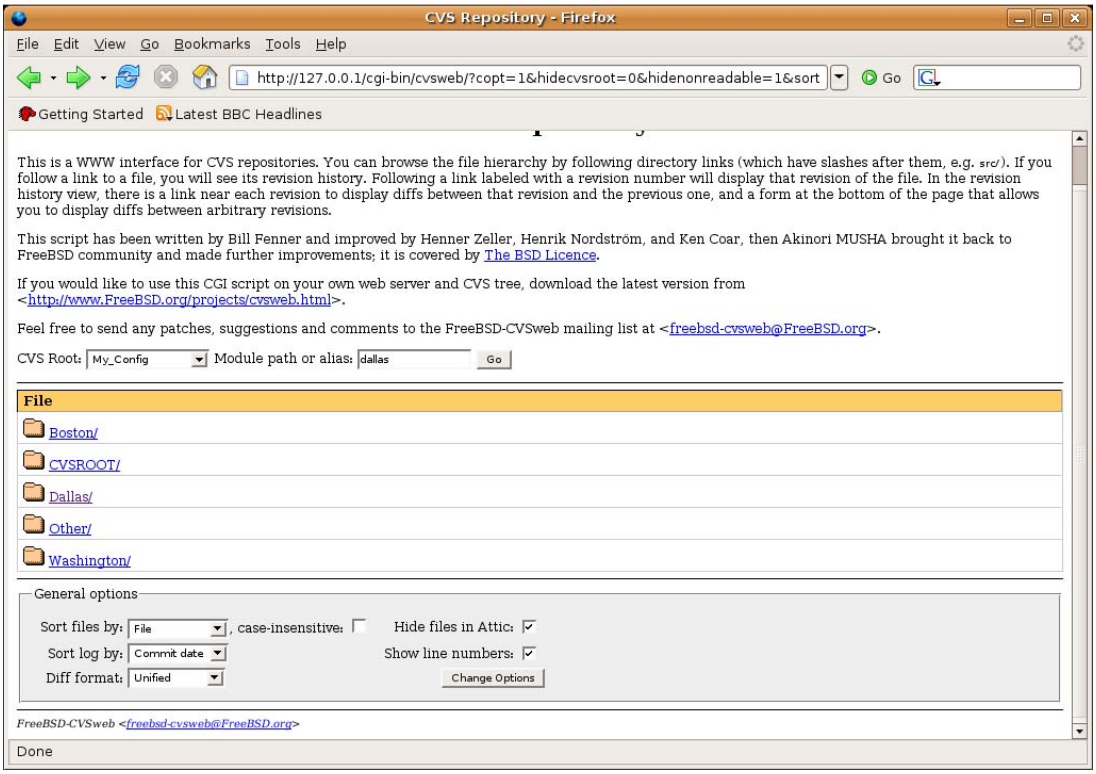

Figura 3.7: Exemplo da interface do Rancid

3.3 Conclusão

As aplicações anteriormente apresentadas podem ser utilizadas separadamente, mas desta forma, não cobrem totalmente o escopo de monitoramento considerado relevante para uma boa administração da rede. Por este motivo, é proposto que estas ferramentas sejam utilizadas de forma complementar, podendo assim, cada qual atuar em sua área de expertise e em conjunto, dispor a equipe de monitoramento um melhor resultado.

Estas aplicações se sobrepõem em algumas áreas, caso não seja possível a utilização de todas elas, o uso de extensões em uma aplicação pode vir a suprir a deficiência pela falta de outra aplicação. Diminuindo a qualidade do resultado obtido, mas mantendo ainda que não tão eficientemente o escopo do monitoramento.

Este último pode ser descrito como a existência dos seguintes fatores:

1. Sistema de alarme de falhas
2. Mapeamento de trafego atual e passado
3. Controle de qualidade dos serviços prestados
4. Histórico de eventos

Estes quatro pontos podem ser cobertos pelo uso das ferramentas anteriormente apresentadas, com o Cacti atuando na modelagem e mapeamento do trafego atual e seu histórico, com um atraso de poucos minutos. Caso a necessidade desta informação exija que tal monitoramento seja praticamente instantâneo, outras aplicações podem entrar nesta área, como por exemplo o NetFlow.

Cobrindo os outros pontos se tem o Nagios, responsável por todo o sistema de alarme e histórico do mesmo, além do tratamento de *Traps*, que são informações geradas pelos equipamentos e acusam alguma mudança no seu comportamento, como uma ativação de interface ou mudança de configuração.

Para finalizar, tem-se o Smokeping que atua prioritariamente na realização de testes de desempenho e no armazenamento destas informações para referências futuras.

O Rancid foi apresentado por sua praticidade em resolver alguns problemas comuns em redes de grandes proporções, como o backup de configurações, a atualização de configurações repetidas em múltiplos equipamentos e um sistema interno que acusa ao gerente caso alguma mudança seja feita manualmente em algum de seus equipamentos.

Capítulo 4

Infraestrutura de Rede Atual da UFF

Apesar da grande abrangência geográfica da UFF, este trabalho tem como foco principal a estrutura localizada fisicamente na região de Niterói (RJ), não contemplando em detalhes as ligações com outras localidades mais remotas da faculdade. Em sua maioria, estas localidades são conectadas logicamente ao equipamento principal da STI através de conexões seguras feitas pela Internet.

4.1 Topologia Física

A estrutura principal são as ligações ópticas entre a STI (Campus do Valonguinho) com a Rede Rio (*Internet Service Provider* (ISP)) e com os outros campi. Este anel principal é denominado como núcleo da rede e cada nó deste núcleo possui a partir de si um foco em estrela que serve cada estrutura de um campus. Os equipamentos que compõem as pontas desta estrela são denominados "Capilares", sendo o primeiro nível de divisão em função do número de usuários. Na hierarquia, ainda existem equipamentos que fazem o acesso do usuário. Contudo, tais equipamentos nem sempre possuem gerência e, por falta de controle e fiscalização, muitas vezes são ligados à rede por um *hub*.

O campus Valonguinho funciona como um centralizador de todo o tráfego direcionado a rede externa, além de ser o local que aloca fisicamente grande parte dos serviços prestados internamente, como mostrado na Figura 4.1.

Como visto na Figura 4.1, o anel que uma vez formava uma estrutura redundante e confiável já não está em operação. Problemas com a fornecedora de energia elétrica e o remanejamento de postes causaram o rompimento de enlaces e a não recuperação imediata dos mesmos, o que causou um acúmulo de falhas, agravando a situação.

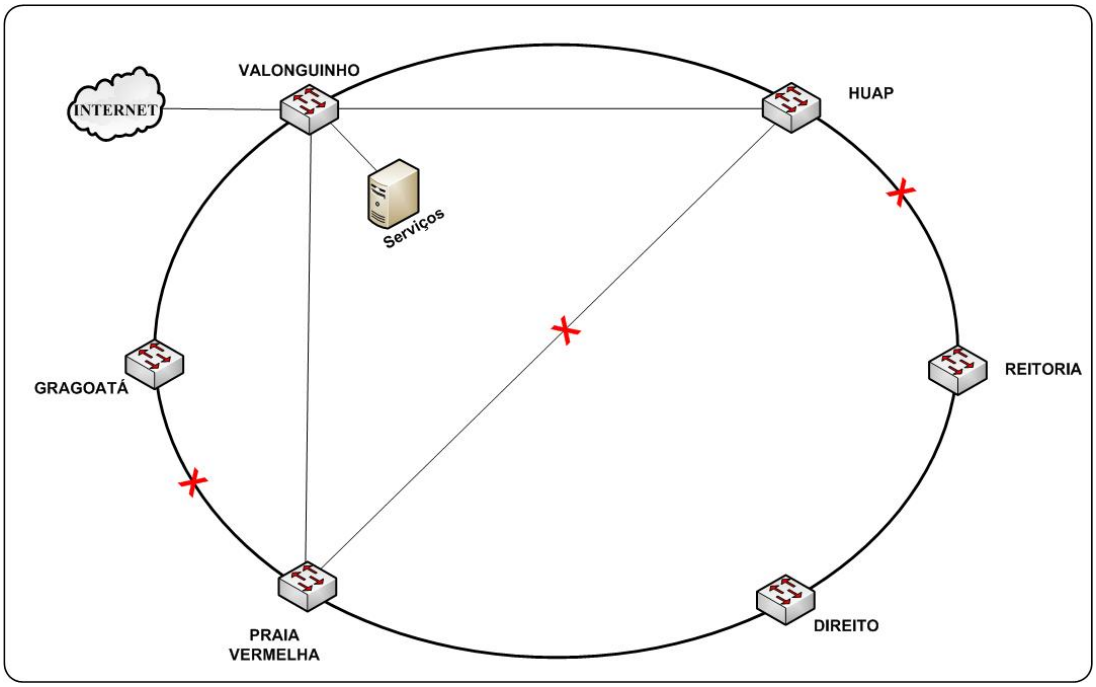

Figura 4.1: Anel da Rede UFF

Demandas de usuários prioritários que se localizavam no final de uma estrutura em cascata e sofriam reflexos de problemas de infraestrutura, principalmente elétrica, em outros campi acabaram por determinar a ligação direta de cada nó diretamente a STI de acordo com a Figura 4.2. Esta ligação foi realizada através de um "jumpeamento"[1] entre as fibras já existentes.

Com o deterioramento causado pelo tempo nos equipamentos principais da rede, atualizações lógicas no seu funcionamento não eram uma realidade. O mau funcionamento destes causava problemas imprevisíveis que não possuíam causa aparente e a solução muitas vezes era a reinicialização dos equipamentos.

4.1.1 Equipamentos

De acordo com a cronologia do desenvolvimento da rede, as licitações mais antigas foram servidas com equipamentos D-Link, pois apresentavam um menor custo, apesar de não serem tão robustos e confiáveis já naquela época. Com o passar do tempo, novas licitações foram feitas, onde a atenção por equipamentos de qualidade melhor esteve presente, começando assim a migração dos equipamentos para um novo fornecedor (3com). Como a necessidade de expansão era urgente, os equipamentos novos foram utilizados

[1]Ligação direta entre portas de Distribuidores Internos Ópticos, sem a conexão a um equipamento ativo.

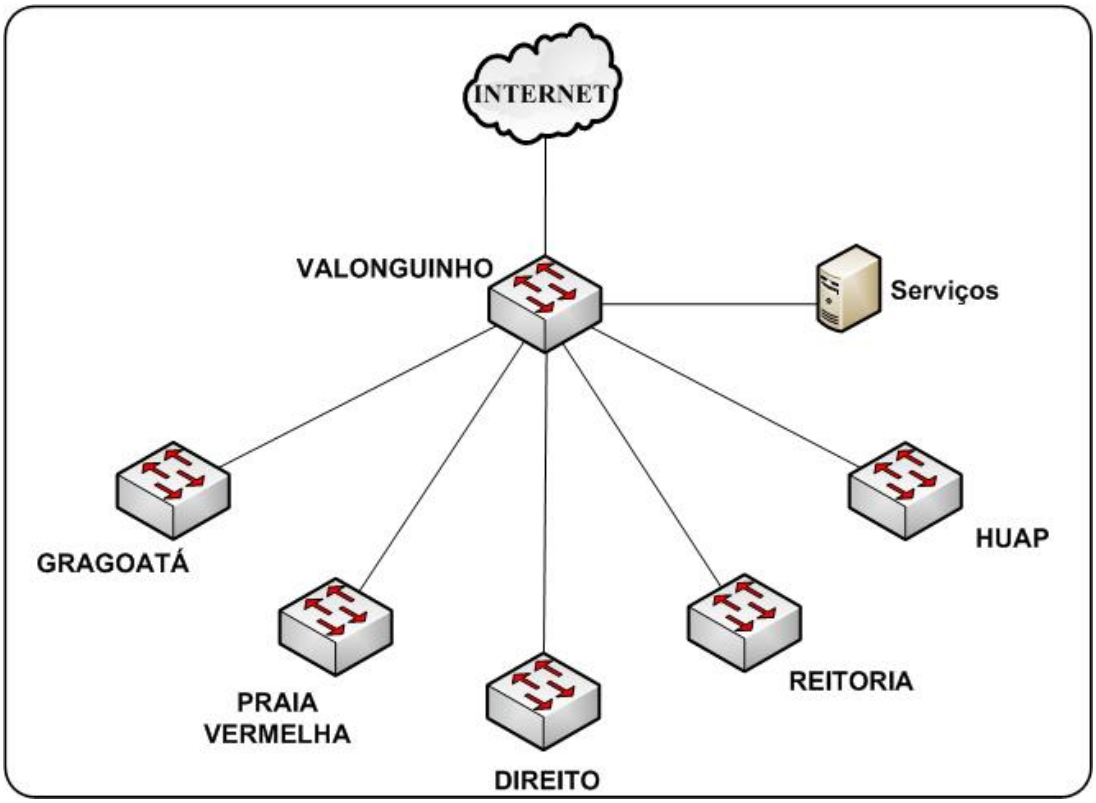

Figura 4.2: Topologia estrela da Rede UFF

para suprir necessidades na capilaridade e distribuição da rede, mantendo o anel central de forma integral ainda com os equipamentos mais antigos.

Esta escolha acabou resultando em alguns problemas já explicitados neste trabalho, pois a extensão da utilização destes equipamentos acarretou problemas causados pelo seu sucateamento.

Estes são os equipamentos mais encontrados na estrutura do núcleo, e na sua capilaridade. Note que grande parte destes equipamentos já não faz parte das linhas de venda dos fornecedores. Maiores informações sobre eles podem ser encontradas nos sites dos fabricantes

- D-Link
  - DXS-3326GSR
  - DGS-3312SR
  - DGS-3324SRi
- 3Com
  - 4500

- 3226
- 2024
- 3300
- 3C16735

Em relação aos equipamentos utilizados no anel central, vale notar que não são equipamentos projetados com esta finalidade, apesar de atuarem como tal. Esta diferença apesar de não parecer grande, acaba pesando em situações de estresse da rede e ou falha generalizada, que resulta na sobrecarga de um nó. Tem-se também uma resiliência reduzida pois não existem medidas de proteção redundante do próprio equipamento, como múltiplas fontes de energia e a sua organização em *stack*.

4.1.2 Problemas Observados

Sobrecarga dos *links* nos *switches* de distribuição

Com o crescimento, a adição de pontos em locais que previamente não estavam preparados acabou causando a sobrecarga de alguns *links* de acesso, causando, assim, alguns gargalos de tráfego na rede.

Perda de pacotes causada por problemas de interoperabilidade.

Ao realizar a mudança de parte dos equipamentos, uma grande perda de pacotes foi notada nas interfaces que interconectam fabricantes diferentes.

Fibras ópticas rompidas

Com a perda sendo inevitável, mas a recuperação do enlace não sendo feita, criou-se uma falta de redundância na parte principal da rede.

Falta de monitoramento em tempo real

Por não possuir um monitoramento, a detecção de falhas e a resolução de problemas se dá de forma tardia e ineficiente.

Falta de histórico da rede

O histórico do monitoramento serve como informação de tendências de uso e pode indicar áreas que precisam de mais atenção, pois estão se tornando mais interconectadas. Este histórico é crucial para a confecção de relatórios e embasa pedidos de novos investimentos.

Não existência de qualquer tipo de alarme

Sem um sistema de alarme, a equipe de suporte ao usuário se encontra alienada de informações sobre a integridade da rede, então todo e qualquer problema reportado por clientes geram uma demanda ao técnico local, muitas vezes desnecessariamente, pois o problema se encontra em outra área a cima na hierarquia

4.2 Topologia Lógica

Logicamente, todo o roteamento externo se concentra em um único nó, que exerce múltiplas funções. Além das funções de roteamento, o mesmo equipamento exerce a função de *firewall* e em alguns momentos, *Intrusion Prevention System* (IPS), que pode causar uma sobrecarga no mesmo.

Internamente, cada nó principal de um campus, é responsável pelas redes presentes no mesmo. Existe uma separação interna feita através do uso de VLANs e todos os usuários recebem IPs públicos através de um serviço de *Dynamic Host Configuration Protocol* (DHCP) centralizado, que é acessível pelas redes por extensões da cada VLAN até o servidor.

As rotas internas são propagadas entre campi pelo protocolo RIPv2, mas por problemas no funcionamento normal dos equipamentos, as rotas nem sempre são obedecidas e/ou propagadas. Isso causa a demanda de configuração e reconfiguração de rotas de forma estática.

Além de prover o acesso à internet, a STI hospeda servidores com diferentes utilidades, tais servidores não possuem uma rede segregada da de usuários, o que pode se tornar uma grande falha de segurança

Capítulo 5

Proposta do Núcleo da Rede UFF

Esse trabalho propõe um protocolo que opera sobre uma topologia de rede em anel, visando com isso resolver alguns problemas da rede atual. São abordados tanto problemas lógicos como problemas físicos.

Uma das maiores limitações que eram impostas ao projeto, é a passagem de novas fibras. Foi definido, anteriormente, pela STI que, por questões orçamentárias, não seria possível o lançamento de novas fibras e/ou ter caminhos físicos diferentes das mesmas. A proposta está focada em utilizar as fibras atualmente disponíveis, recuperando enlaces indisponíveis e mudar os protocolos utilizados para o funcionamento do núcleo.

A atualização dos equipamentos principais da rede também faz parte do escopo deste projeto, atualização esta que será feita com equipamentos licitados pela UFF e que são mais elaboradamente descritos na seção 5.2 .

Foram feitos experimentos com o protocolo RRPP para verificar se ele de fato serviria para o núcleo da Rede IP da UFF.

5.1 Topologia

A topologia proposta da rede é um grande anel que liga os 6 campi da UFF.

Como pode ser visto nas Figura 5.1 e Figura 5.2, cada campus tem dois *switches* configurados em *stack*. Os *switches* configurados em *stack*, são ligados por um cabo de 10 Gb para atender a capacidade do núcleo da rede UFF que também está ligado à 10 Gb entre cada campus. Cada *switch* capilar recebe uma fibra de cada *switch* em *stack* para a redundância de *link* e agregação.

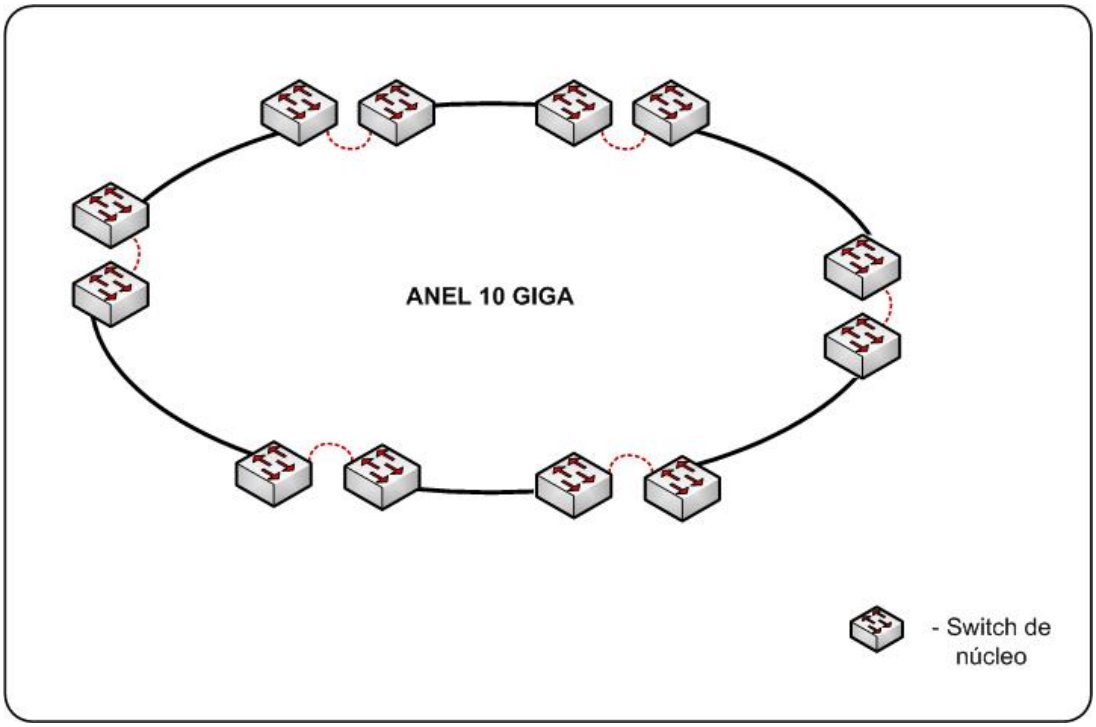

Figura 5.1: Anel 10 GIGA da Rede UFF

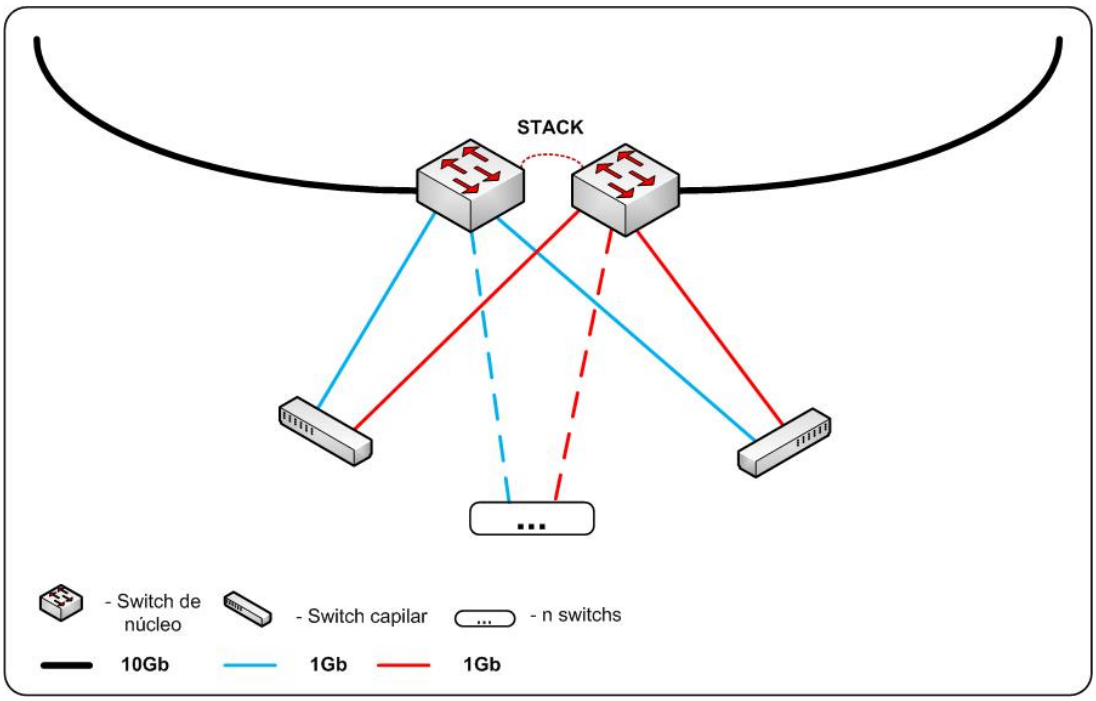

Figura 5.2: Ligação dos Capilares da Rede UFF

Essa configuração em *stack* faz com que os dois *switches* se comportem como somente um, ou seja, um *switch* com o dobro de portas, mas a mesma capacidade computacional. Com isso, caso um *switch* queime ou desligue, o outro continua ativo e com as mesmas configurações, existe uma redundância em nível de equipamento.

Dessa forma pode ocorrer um rompimento de fibra em qualquer ponto do anel que todos os equipamentos ainda possuem um caminho físico para se comunicarem. Provendo assim uma tolerância a falhas. Como também é possível que qualquer *switch* do núcleo queime ou desligue, através do *stack* e seus *links* agregados, os capilares ainda teriam um caminho válido a percorrer.

O RRPP enxerga a falha em um dos *switches* em *stack* como um rompimento do anel. Por isso e por causa da redundância dos *links*, os capilares podem continuar a se comunicar.

5.2 Equipamentos Utilizados

Os equipamentos para a realizar a função proposta foram adquiridos pela STI para uma melhoria da rede infraestrutura da UFF. São equipamentos da H3C, empresa que surgiu com a compra da 3com pela HP. Durante a licitação dos equipamentos, foram descritos equipamentos para o core da rede e sua capilaridade de primeiro nível. Os equipamentos entregues pela empresa ganhadora da licitação foram:

- HP A5500-24G-SFP EI
- A5500-24G EI

Os equipamentos escolhidos para o núcleo são *Switches Layer 3*[1], indicados para a utilização em empresas ou instituições de médio e pequeno porte, a possibilidade da utilização de *Small Form-factor Pluggables*[2] (SFP) é de grande valia pois a capilaridade da rede se da de forma integral através de enlaces de fibra óptica. A versatilidade dos equipamentos no que condiz a novos módulos que possam adequá-los a novos cenários futuros também é importante, pois novos transceptores[3] com a capacidade de manter enlaces com a taxa de transferência na casa dos 10 Gb já está sendo implementados.

Com o crescimento da rede, a possibilidade de segregar ainda mais a forma que o roteamento entre sub-redes é tratado se torna cada vez mais atraente, então o suporte que estes novos equipamentos possuem ao OSPF os torna ainda mais uteis em suas funções e possiveis futuras atribuições.

[1]Equipamento de comutação capaz de realizar algumas funções da camada de redes

[2]Interface *plug-and-play* que pode ser utilizada para disponibilizar uma diferente gama de conectores

[3]Dispositivo que combina um transmissor e um receptor utilizando componentes de circuito comuns para ambas funções num só aparelho

5.3 Problemas Observados

Um dos problemas dessa arquitetura é que ela suporta apenas uma falha para todo o anel. Não chega a ser um grande problema, mas é importante ressaltar que, em um pico de energia, é possível que dois ou mais equipamentos de rede fiquem inutilizados temporária ou permanentemente. Nesse caso, a rede não teria um caminho alternativo e estaria fora do ar.

É possível também que, caso aconteça um rompimento de fibra, o tempo necessário para repará-la seja grande. Nesse caso, a rede ficaria sem proteção contra falhas durante todo o período até o reparo ser realizado. Mais uma vez, um bom sistema de monitoramento e de alertas se faz necessário, pois caso aconteça alguma falha, o sistema fica sem proteção alguma.

5.4 Comparação

Ambos cenários são topologicamente idênticos, diferindo apenas no reestabelecimento de *links* problemáticos e na readequação da infraestrutura elétrica que existe nos pontos de instalação de equipamentos. Entretanto, as mudanças na utilização de protocolos de proteção do anel traz um ganho na velocidade de resposta a falha, que deixa de ser manual com um tempo que poderia ser de poucos minutos a algumas horas, dependendo do nó afetado, e passa a ser uma resposta em milissegundos feita de forma automática.

Como essa nova realidade traz com si a necessidade do monitoramento, outros ganhos podem ser frutos dessa necessidade imposta pela mudança, com ele outras possibilidades se tornam possível para melhor prover o serviço de acesso à internet e a qualidade nas aplicações geridas pela STI.

Capítulo 6

Experimentos com o RRPP

Esse capítulo tem como objetivo apresentar os experimentos feitos para avaliar o desempenho do protocolo RRPP em um anel, simulando um cenário similar ao do núcleo da rede da UFF. Os equipamentos utilizados para a realização destes experimentos foram emprestados pela Superintendência de Tecnologia da Informação da Universidade Federal Fluminense e foram realizados os seguintes experimentos:

Quantitativos:

- Núcleo RRPP,
  - Funcionamento em uma situação normal.
  - Tempo de resposta ao rompimento do anel em um único ponto.
  - Tempo de resposta ao reestabelecimento do anel no ponto. rompido

Estes experimentos foram realizados conforme topologia mostrada na Figura 6.1

Qualitativos:

- Ligação Skype [7], percepção de desempenho de uma ligação em meio ao rompimento/restabelecimento do anel.
- Download de um arquivo via HTTP durante o rompimento/restabelecimento do anel.

Os experimentos contaram com seis *switches* H3C, que serão futuramente utilizados no projeto e dois notebooks. As conexões entre os *switches* foram feitas com cabos de fibra óptica monomodo e as conexões entre os *switches* e os notebooks foram feitas com

cabos UTP CAT5e, ambos provisionados pela STI. Os equipamentos estão descritos ao final desse capítulo.

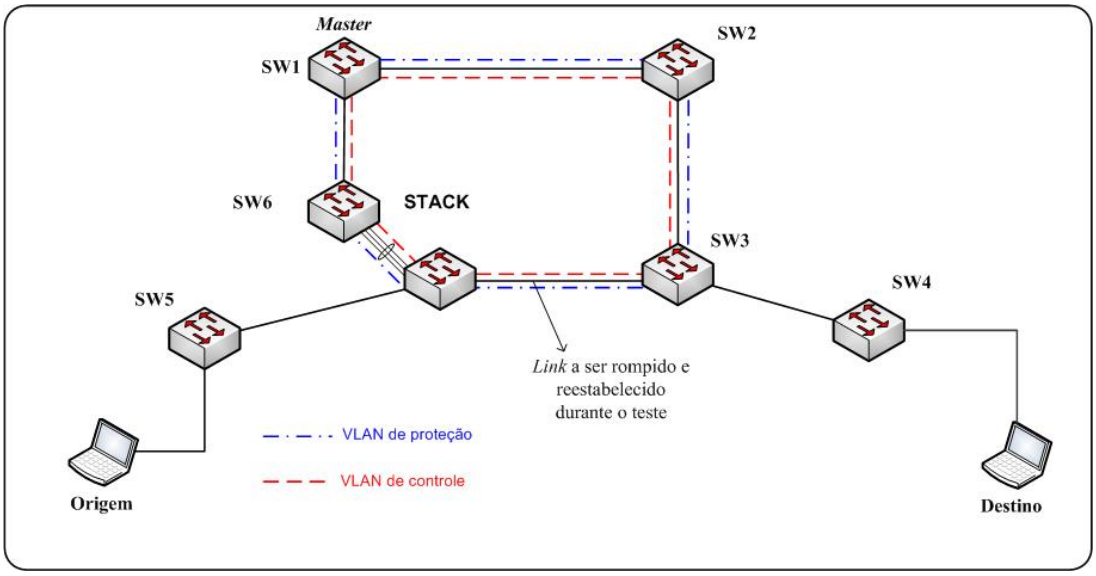

Figura 6.1: Topologia do experimento ping

6.1 Metodologia dos Experimentos

O objetivo principal foi verificar o tempo que o RRPP demora para reestabelecer o *link*, ou seja, por quanto tempo não há conectividade entre as máquinas que estão conectadas por um caminho que está passando pela fibra que foi rompida. As ferramentas utilizadas foram:

- PING - ferramenta que utiliza o protocolo ICMP. [8]
- Iperf - ferramenta para medir a capacidade do canal na rede. [9]
- tcpdump - ferramenta para captura de tráfego na placa de rede. [10]
- Wireshark - ferramenta para análise de protocolos de rede. [11]
- Apache - servidor web. [12]
- Chrome - navegador web. [13]
- udpflood - *script* na linguagem perl para geração de pacotes UDP. [14]
- Skype - ferramenta para efetuar ligações pela internet.

6.1.1 Metodologia do Experimento com a ferramenta PING

Como o objetivo do experimento envolvia a medição do tempo de indisponibilidade causado pela falha, a ferramenta ping, que proporciona o tempo decorrido entre o envio de um pacote e o de chegada da resposta do destino, pareceu ser a melhor forma de obtê-lo. Esta ferramenta possui muitas formas de customização deste envio e poderia ser adequada as necessidades do experimento.

Sendo assim, a ferramenta PING foi utilizada para gerar e enviar pacotes durante 10 segundos com um intervalo entre pacotes de 0.001 segundos (1 ms). Para tal, utilizou-se o comando:

```
ping -i 0.001 -w 10 -D  [IP de destino]
```

-i : intervalo de espera em segundos do envio de cada pacote
-w : tempo de execução total do comando
-D : imprimi na tela a hora em que o *reply* é recebido (unixtime[1] + microsegundos)

O resultado deste comando é da forma que se segue:

```
[1386792922.523867] 64 bytes from 192.168.99.17: icmp_seq=5 ttl=64
time=0.295 ms
```

Com o intuito de obter medidas de forma mais precisa, outra ferramenta, tcpdump, foi utilizada para obter dados sobre o tempo em que ocorria a chegada do pacote na interface de destino, podendo assim remover o tempo de processamento e resposta que este computador poderia inserir no resultado final. A análise dos dados foi feita com o auxílio do Wireshark.

```
tcpdump -i eth0 -s 65535 -w ping1-t$i.pcap
```

[1]Número em segundos passados desde primeiro de Janeiro de 1970, sem considerar anos bissextos

-i : Interface desejada
-s : tamanho de bytes a ser capturado de cada pacote
-w : direcionamento da saída do comando para um arquivo
$i : número da iteração

Vale notar que o tcpdump necessita de alguns segundos para inicialização completa. Sendo assim, ele precisa ser iniciado antes da geração do tráfego e rompimento do anel. Para isso, utilizou-se um atraso de três segundos entre sua execução e o incio da transmissão de pacotes.

Para obter um resultado estatístico válido, os experimentos foram realizados múltiplas vezes. Um *script* foi elaborado para automatizar essa repetição, diminuindo, assim, a possibilidade de erros gerados por eventos isolados.

O *script* basicamente executa 20 vezes os seguintes passos:

1. Executa tcpdump
2. Espera três segundos
3. Executa o ping com duração de dez segundos
4. Espera três segundos
5. Termina o tcpdump

Durante os experimentos, observou-se que a ferramenta ping, quando não recebe um *reply*, espera 10 ms antes de enviar o próximo ICMP e esse comportamento se repete até que um *reply* seja recebido. Após o recebimento do *reply*, a ferramenta volta a enviar pacotes de 1 ms em 1 ms.

Para verificar o tempo em que as máquinas ficam sem conexão, ou seja, o tempo entre o rompimento do anel até o estabelecimento de um novo caminho, foi utilizada a aproximação representada pelo maior tempo entre a diferença de chegada dos *replys* do ping *flood*.

6.1.2 Metodologia do experimento com a ferramenta UDP *Flood*

Após a análise dos experimentos realizados com a ferramenta ping, um possível ponto de falha nas medições foi identificado, pois o comportamento desta ferramenta frente

a uma situação de falha poderia alterar o tempo final de resposta encontrado. Esta realização motivou a busca por uma outra ferramenta que, sendo mais simples não utiliza confirmações de recebimento ou a identificação de sequência de pacotes. Assim a utilização de um *script* que funcione com o protocolo UDP e enxurrada de pacotes se mostrou uma melhor solução.

Com a ferramenta em perl para fazer UDP *flood*, foi feito o *flood* de uma máquina para a outra. A máquina destino dos pacotes UDPs capturava os pacotes com a ferramenta tcpdump.

A configuração do tcpdump foi:

```
tcpdump udp -w udpflood-up-t$i.pcap
```

udp : filtrar somente pacotes udp
-w : direcionamento da saída do comando para um arquivo
$i : número da iteração

Novamente, foi utilizado um *script* para auxiliar no experimento. O *script* executa 50 vezes os seguintes passos:

1. Executa tcpdump
2. Espera três segundos
3. Imprime na tela para desconectar ou conectar a fibra
4. Espera três segundos
5. Termina o tcpdump

Com base nos arquivos de captura de pacotes, é possível verificar os momentos de chegada. O experimento visa descobrir quanto tempo as máquinas ficaram sem conectividade. Para tal, calcula-se a diferença entre os tempos de chegada dos pacotes.

Para calcular a diferença entre os tempos de chegada dos pacotes, o arquivo de captura de pacotes foi analisado com o seguinte comando:

```
tcpdump -ttt -r udpflood-down-t$i.pcap > udpflood-down-difftime-t$i.txt
```

-ttt : Imprime a diferença de tempo em microssegundos entre a linha atual e a anterior do arquivo de *dump*.
-r : Lê os pacotes de um arquivo criado com a opção -w"
> : Redireciona a saída para um arquivo
$i : número da iteração

A Linha gerada é:

```
00:00:00.000027 IP 192.168.99.17.49514 > archlinux-laptop.7785:UDP,
length 862
```

Onde o atraso entre a recepção de pacotes é o primeiro campo. Para obter o tempo em que não existia conectividade, basta verificar o maior tempo dentre todas as linhas do arquivo.

6.2 RRPP Núcleo - Experimentos Quantitativos

Este experimento verifica como a rede se comporta quando configurada com RRPP no núcleo. Os experimentos para esse cenário contém três etapas:

- Anel íntegro - anel completo e sem falhas.
- Quebra do anel - onde apenas um cabo do anel é rompido.
- Reestabelecimento do anel - onde o cabo rompido no experimento de quebra de anel é reestabelecido.

Foram realizados dois experimentos para verificar o tempo de reestabelecimento de conectividade após o rompimento e o restabelecimento do anel. Os experimentos foram feitos de acordo com a metodologia explicada acima.

6.2.1 Anel Íntegro

Para se ter uma referência do funcionamento em condições normais, foi realizado o experimento sem que nenhuma alteração no anel ocorresse. Os resultados obtidos demonstram que nenhum pacote foi perdido em nenhuma das 20 iterações feitas com o experimento do ping. Não foram realizados experimentos com udpflood nesse cenário.

6.2.2 Rompimento do anel

Uma simulação de rompimento foi feita em uma das fibras pela qual o tráfego passava. A desconexão de uma de suas pontas simula esse efeito.

6.2.2.1 PING

Para o experimento do ping, a desconexão da fibra ocorre dentro do intervalo de 10 segundos no qual o ping está sendo executado.

Para melhor visualização dos resultados do experimento com o ping, tem-se a Figura 6.2 e a Tabela 6.1

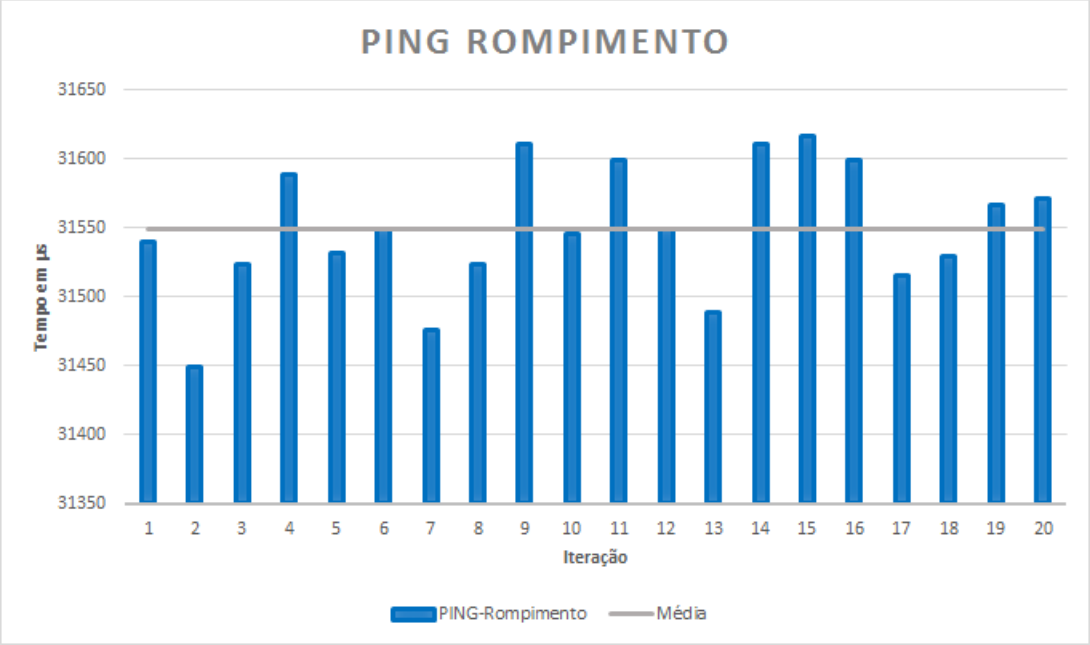

Figura 6.2: Resultado do experimento ping durante o rompimento

As conclusões retiradas a partir destes resultados e sua análise se estão expostas na seção 6.2.4.

Tabela 6.1: Resultado do experimento ping durante o rompimento

Parâmetro	**Valor μs**
Mínimo	31449,00
Máximo	31616,00
Média	31549,00
Desvio padrão	46,77

6.2.2.2 UDP *Flood*

Para o experimento do udpflood, o cabo é desconectado na hora em que o *script* imprime na tela para remover. Essa ocasião é 3 segundos após o tcpdump ter iniciado. Portando o tcpdump está capturando pacotes quando ocorre o rompimento do anel. Esses passos ocorrem 50 vezes, assim se tem um melhor resultado estatístico

Para melhor visualização dos resultados do experimento udpflood, tem-se a Figura 6.3 e a tabela 6.2

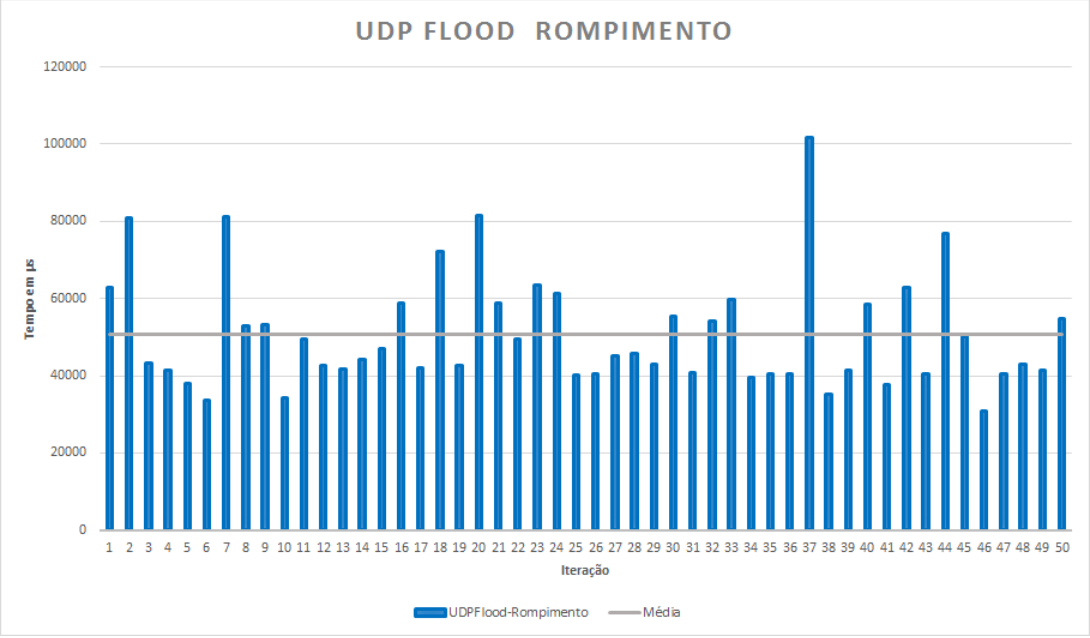

Figura 6.3: Resultado do experimento udpflood durante o rompimento

Tabela 6.2: Resultado do udpflood durante o rompimento

Parâmetro	**Valor μs**
Mínimo	30808,00
Máximo	101599,00
Média	50727,30
Desvio padrão	14733,35

As conclusões retiradas a partir destes resultados e sua análise se estão expostas na seção 6.2.4.

6.2.3 Reestabelecimento do anel

De forma similar ao experimento anterior, tem-se agora a simulação de uma fibra sendo fundida, com seu pleno funcionamento restabelecido. O experimento começa com o anel já com uma fibra desconectada. A conexão da ponta desconectada da fibra provoca o efeito de uma fibra sendo fundida. Essa conexão foi realizada em um tempo qualquer durante a execução dos experimentos.

6.2.3.1 PING

Assim como no experimento de rompimento da fibra, a reconexão é feita durante a execução do comando ping. Isso ocorre 20 vezes. Para melhor visualização dos resultados do experimento com o ping, tem-se a Figura 6.4 e a tabela 6.3

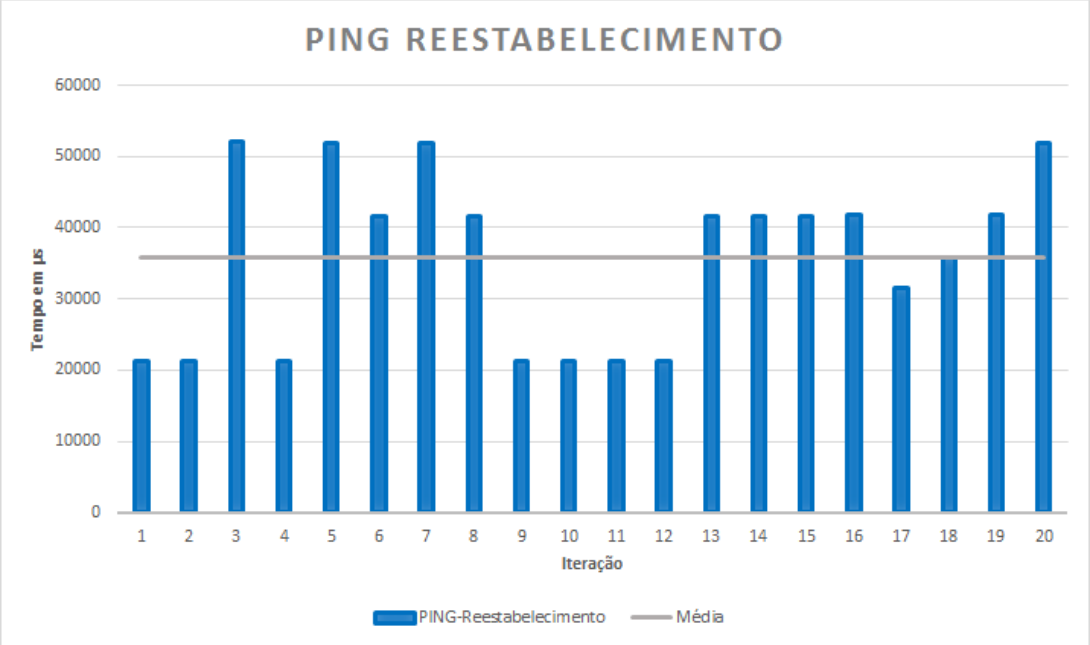

Figura 6.4: Resultado do experimento ping durante o reestabelecimento

As conclusões retiradas a partir destes resultados e sua análise se estão expostas na seção 6.2.4.

Tabela 6.3: Resultado do experimento ping durante o restabelecimento

Parâmetro	**Valor μs**
Mínimo	21249,00
Máximo	52021,00
Média	35784,21
Desvio padrão	12066,46

6.2.3.2 UDP *Flood*

Da mesma forma que no caso do rompimento do anel para o experimento do udpflood, o anel é conectado durante a captura dos pacotes com o programa tcpdump. Isso ocorre 50 vezes.

Para melhor visualização dos resultados do experimento udpflood, tem-se a Figura 6.5 e a tabela 6.4

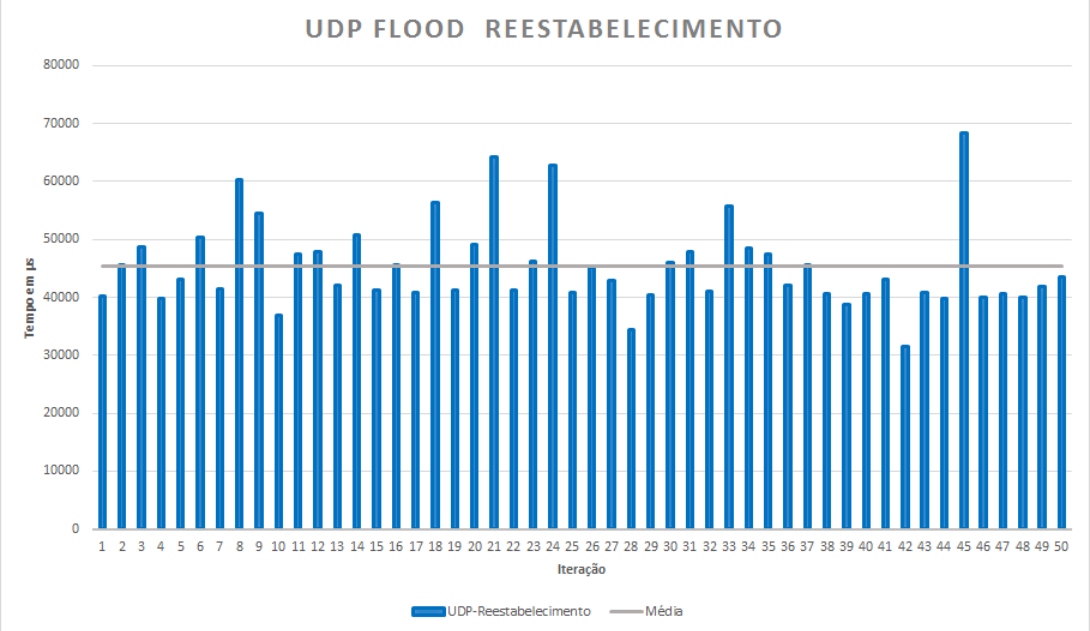

Figura 6.5: Resultado do experimento udpflood durante o reestabelecimento

Tabela 6.4: Resultado do experimento udpflood durante o reestabelecimento

Parâmetro	**Valor μs**
Mínimo	31562,00
Máximo	68450,00
Média	45316,56
Desvio padrão	7428,13

As conclusões retiradas a partir destes resultados e sua análise se estão expostas na

seção seguinte.

6.2.4 Conclusões Experimentos Quantitativos

Com base nos experimentos, é possível observar que, com o RRPP ativado no core da rede, o rompimento ou o reestabelecimento de uma fibra deixa a rede sem conexão por algumas dezenas de milissegundos. Isso atende aos requisitos do core de rede UFF estabelecidos pela STI.

O experimento com a ferramenta udpflood foi realizado devido ao comportamento do PING, o qual deixa de enviar as mensagens no intervalo especificado e sim com um intervalo de 10 ms quando a perda de um pacote ocorre. Buscando uma maior consistência no intervalo de envio de pacotes, esta ferramenta foi priorizada.

Os experimentos com o PING e com o udpflood apresentam médias diferentes. Isso se explica porque o atraso medido apresenta grande variabilidade, como mostrado na Figura 6.6. Repetições do mesmo experimento mostraram resultados no qual a média do udpflood foi inferior à do ping, mas, novamente, não é possível dizer que um foi menor que o outro, considerando-se a barra de erro. Acredita-se que essa variabilidade pode ter ocorrido por fatores externos, como o tempo para o restabelecimento das tabelas ARP, processamento interno dos *switches*, entre outros.

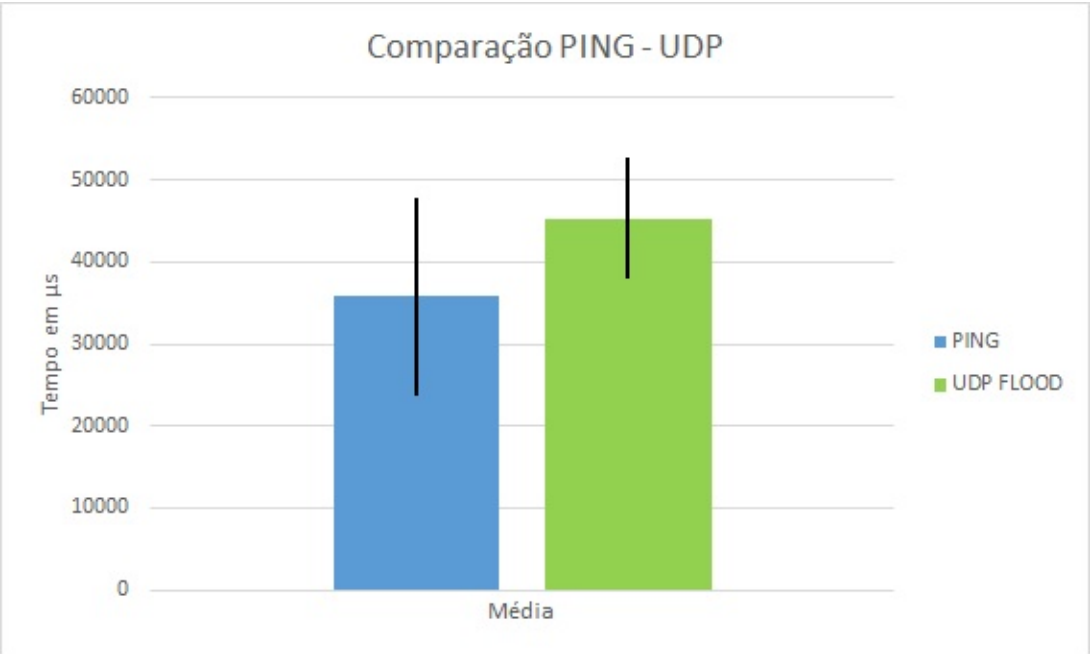

Figura 6.6: Comparação dos resultados dos experimentos ping e udpflood

É importante notar que o objetivo do experimento não era medir com precisão o tempo

de reestabelecimento da conexão, ou seja, verificar se o RRPP realmente converge em no máximo 50 ms, e sim se esse tempo estaria de acordo com os requisitos pré estabelecidos. A transparência da falha no anel para o usuário é o fator mais importante no ponto de vista da STI. Logo, verificar se o usuário sofreria algum impacto foi o objetivo do experimento. Como será visto nos experimentos qualitativos mostrados a seguir, não houve percepção do usuário.

Com os experimentos feitos é concluído que o protocolo RRPP é de fato uma boa escolha na arquitetura de proteção de *loop* no core da rede.

6.3 Experimentos Qualitativos

Com a intenção de mensurar qualitativamente estes resultados, foram feitos dois outros experimentos que reflitam situações cotidianas de uso dos usuários finais.

6.3.1 Ligação Skype

Durante a realização de uma chamada entre dois computadores, um na rede estruturada pelo anel e outro em uma outra rede, conectados pela internet, com a transmissão de voz e imagem, simulamos o rompimento e o reestabelecimento do anel óptico, como mostrado na Figura 6.7.

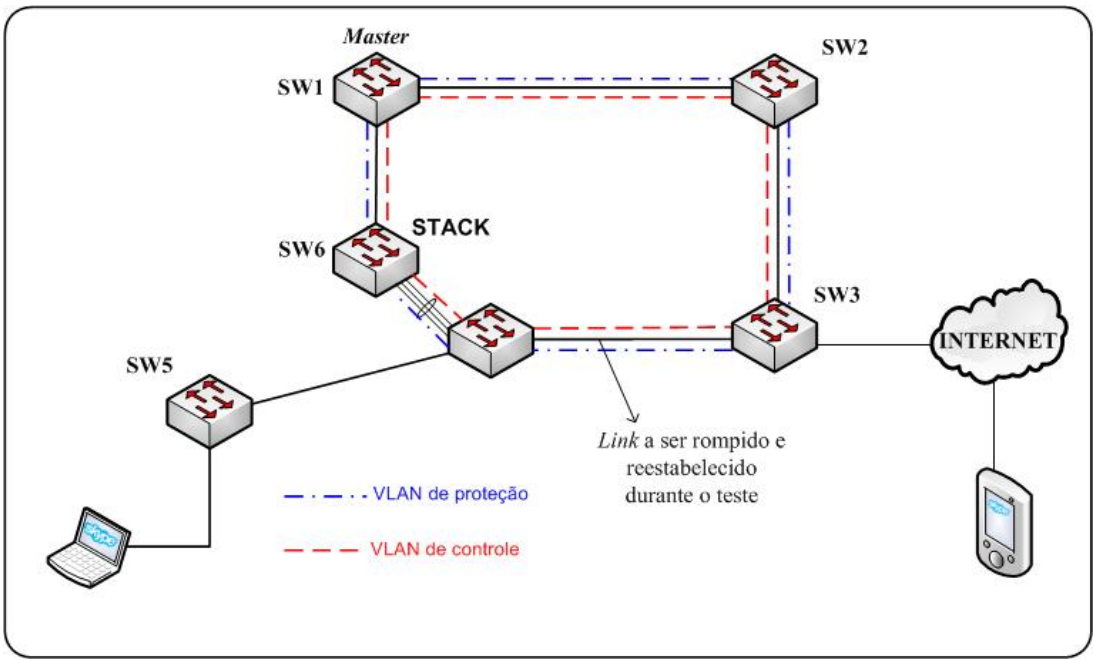

Figura 6.7: Topologia do experimento da Ligação Skype

Ambas ações não geraram impacto perceptível para os usuários. Uma filmagem do experimento foi realizada e está disponível para visualização no site: `https://docs.google.com/file/d/0B5_V7Jqc7nOYLVJxc0dMQmg4SFk`

6.3.2 Download de um arquivo via HTTP

Esse experimento utiliza apenas a rede local, com um computador com um *webserver* servindo o arquivo a ser baixado e outro como cliente, como visto na Figura 6.8. O *link* entre **SW3** e **SW6** é rompido durante o download.

Novamente, nenhuma diferença no desempenho do download do arquivo pode ser notada. Uma filmagem do experimento foi realizada e esta disponível para visualização no site: `https://docs.google.com/file/d/0B5_V7Jqc7nOYdEZ4QzM5NFpVVFE`

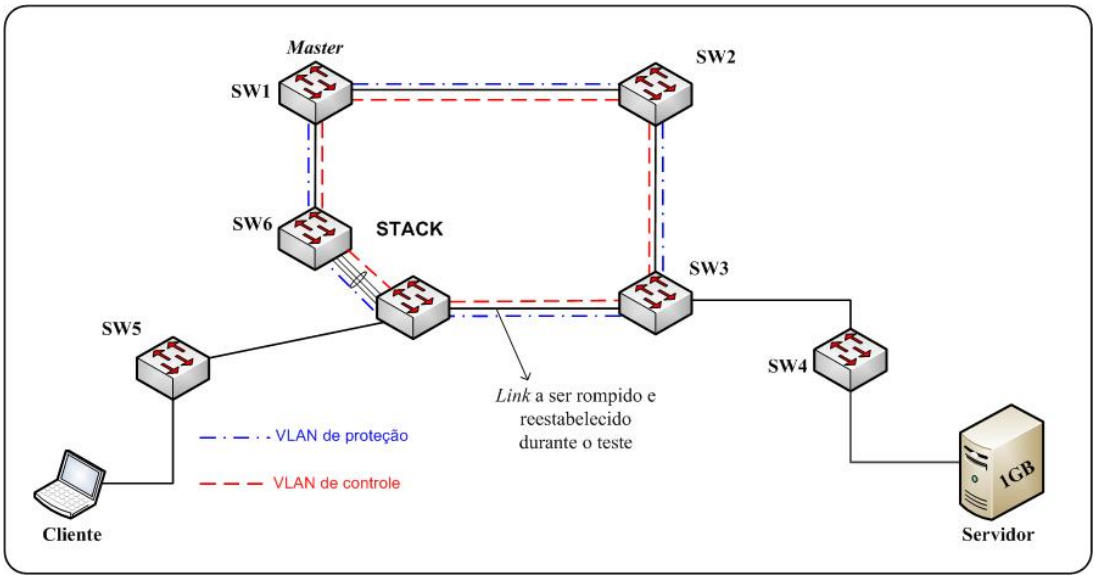

Figura 6.8: Topologia do experimento Download

Para mostrar que realmente o tráfego está passando pelo anel. Na segunda parte do vídeo é desconectada a outra perna do anel. Dessa forma o caminho para o servidor é perdido, e é possível ver o tráfego parando.

6.4 Comportamento do Tráfego

Como o protocolo RRPP utiliza o bloqueio das VLANs de transmissão de dados de uma das interfaces do *master node* como forma de evitar um *loop* de pacotes. Para comprovar o comportamento do tráfego, no que condis com seu trajeto de escolha, foi

utilizada a ferramenta Cacti, que gera uma resposta gráfica para as informações de uso de cada interface. A seguir, estes gráficos serão apresentados, dando o foco em cada nó, para melhor entendimento das informações apresentadas.

Além disso, o programa iperf foi utilizado parar gerar um tráfego fictício por aproximadamente duas horas, facilitando assim a visualização gráfica, após uma hora foi provocado um rompimento do enlace que estava sendo utilizado, provocando uma restruturação das tabelas ARP/ND e assim, um novo fluxo para as informações. Os gráficos apresentados nas próximas seções demostram que após o rompimento todo tráfego começa a passar por outro enlace. Estas informações são expostas utilizando um único *switch* como referência para facilitar o entendimento. A topologia inteira está demonstrada na Figura 6.9

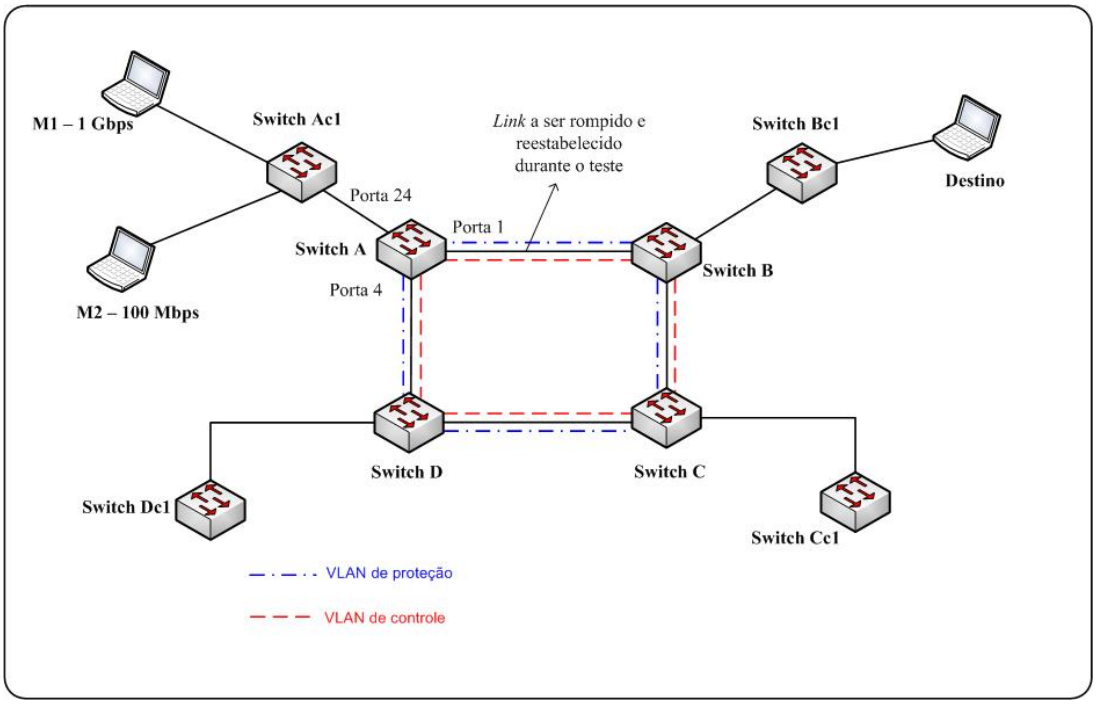

Figura 6.9: Topologia do experimento de comportamento do tráfego

- *Switch* A

Como pode ser visto na Figura 6.9, o *switch* A é a conexão das máquinas M1 e M2 ao anel. As máquinas M1 e M2 estão gerando tráfego, uma a 1 Gbps e a outra a 100 Mbps.

A Figura 6.10 mostra o tráfego na porta 1 do *switch* A. As vinte horas e trinta minutos o anel é rompido, e o tráfego na porta 1 cessa.

Observação: Os outros picos de tráfego as 18:00 (experimento de funcionamento) e 21:00 (experimento de reestabelecimento), são outros experimentos que foram realizados ao mesmo tempo. Eles não devem ser considerados.

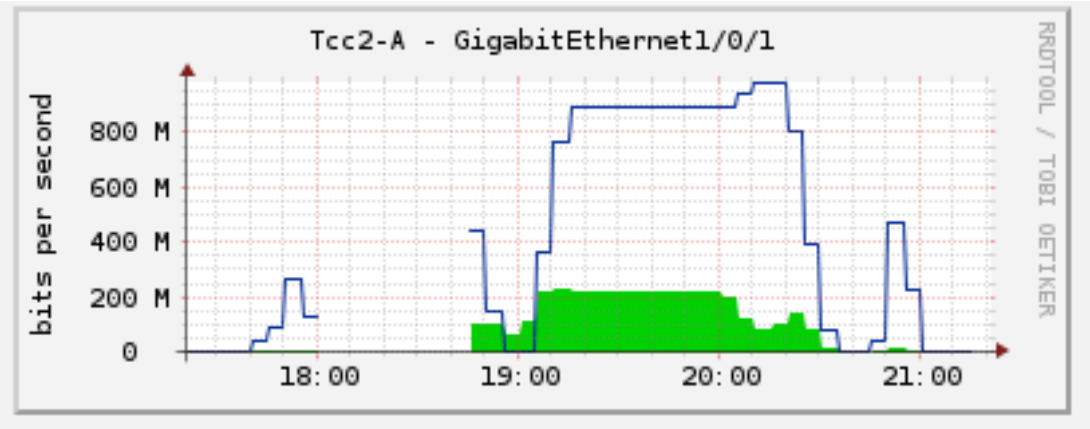

Figura 6.10: Porta 1 do *Switch* A

A Figura 6.11 mostra a porta 4 do Switch A, que é a outra ligação do anel.

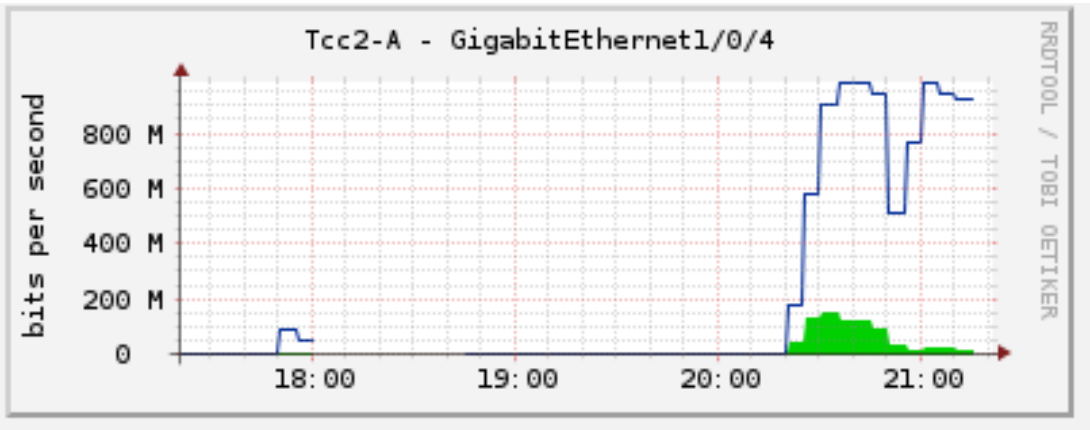

Figura 6.11: Porta 4 do *Switch* A

É fácil notar que antes todo tráfego estava sendo bloqueado pelo RRPP, porém agora que o anel foi rompido, o tráfego foi liberado na porta 4 e todo tráfego gerado pelo iperf começa a chegar por essa porta.

Como a porta 24 não faz parte do anel, ela simula uma interface de um cliente em um *switch* capilar. Para os clientes, o tráfego não deve ser afetado devido ao rompimento do anel. Isso fica claro na Figura 6.12.

6.5 Configuração das máquinas dos Experimentos

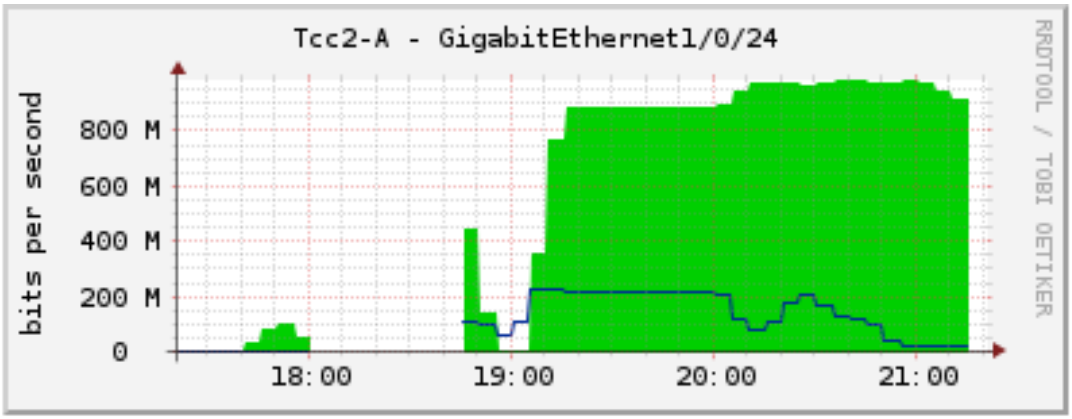

Figura 6.12: Porta 24 do *Switch* A

Tabela 6.5: Máquina 1

Sistema Operacional	Linux sfrique-portable 3.12.3-1-ARCH SMP PREEMPT Wed Dec 4 21:59:02 CET 2013 i686 GNU/Linux
Mémória RAM	4 Gb
Processador	Intel® Core™ i5-3317U CPU @ 1.70GHz × 4
Ultima Atualização	11 de Dezembro de 2013

Tabela 6.6: Máquina 2

Sistema Operacional	Linux sfrique 3.12.3-1-ARCH SMP PREEMPT Wed Dec 4 21:59:02 CET 2013 i686 GNU/Linux
Mémória RAM	2 Gb
Processador	Intel® Core™2 Duo CPU T5450 @ 1.66GHz × 2
Ultima Atualização	11 de Dezembro de 2013

Tabela 6.7: Máquina 3

Sistema Operacional	Windows 8.1
Mémória RAM	4 Gb
Processador	Intel® Core™ i5-3317U CPU @ 1.70GHz × 4
Última Atualização	05 de Dezembro de 2013

Capítulo 7

Conclusão

Conforme desenvolvido no corpo deste trabalho, tendo o crescimento da importância das redes de comutação em mente e a necessidade de uma maior confiança e desempenho na infraestrutura da UFF, um projeto de reestruturação e atualização foi buscado e foi disponibilizado por uma empresa terceira a instituição.

Para melhor analisar o cenário atual e todas as nuances que tal projeto propunha, um estudo teórico mais aprofundado sobre o protocolo proposto e seus pares foi conduzido. Nele, as diferenças foram expostas e sua superioridade aos protocolos da família STP foram notados. Já seu funcionamento, quando comparado ao protocolo similar, mas não proprietário, mostrou que ambos possuem um desempenho similar.

Com a introdução desta nova tecnologia de rede, o RRPP, problemas como a falha de um enlace se tornam ainda mais críticos. Apesar do ganho claro em tempo de resposta, ao se inserir um mecanismo automático de tratamento da falha, ela pode passar despercebida, ficando assim, sem um alerta que de início ao processo de solução definitiva da mesma. Enquanto esta falha não for tratada, a redundância inicialmente imposta à rede como mecanismo de defesa deixa de existir, tornando esta rede mais propensa a uma falha de grandes proporções, com a interrupção do serviço para parte dos usuários.

Além da necessidade exposta anteriormente, a solução de outros tipos de problemas que já existiam, como a sobrecarga de enlaces, a falta de conhecimento sobre o perfil de utilização da rede, indicadores de desempenho, tem muito a ganhar com sistema de monitoramento eficiente.

Apesar da proposta realizada para os softwares de monitoramento ser integralmente livre, existem muitos outros programas que realizam funções semelhantes, desenvolvidos e mantidos por empresas de renome. Ela engloba de uma forma geral todos os parâmetros

necessários, mas a utilização parcial de cada ferramenta de forma efetiva já seria de grande valia. Estas ferramentas precisam estar sempre atualizadas e condizentes com a realidade atual da rede para poderem gerar respostas válidas.

Com base na proposta recebida inicialmente para a rede do STI, um estudo prático mais aprofundado da rede foi realizado. Com ele, novos problemas puderam ser expostos e uma análise critica mais voltada aos problemas que assolam o dia a dia puderam ser conhecidos. Esta análise se dá através do estudo da topologia física e lógica que ela adota.

Após a análise profunda das informações adquiridas anteriormente, uma proposta final da topologia é explicitada com as principais diferenças e seus pontos de vantagem e desvantagem quando comparada a topologia que está implementada.

Após a parte teórica estar finalizada, a comprovação de sua funcionalidade é estudada com uma série de experimentos que envolvem situações pouco prováveis, como a quebra de um enlace, ou o rompimento de múltiplos enlaces simultaneamente. Com estes experimentos, é possível notar que o tempo de convergência da rede para o evento de uma única falha é dentro dos padrões especificados pelas necessidades de projeto.

Testes qualitativos que refletem a experiência de um usuário da rede no momento de uma falha demonstram que esta é praticamente imperceptível para as aplicações de camadas mais altas, sendo esta falha praticamente transparente ao usuário.

Um fato importante que deve ser levado em consideração, é que com base nos experimentos destes trabalho, o projeto que tinha sido recebido pela STI, sofrerá alterações. A arquitetura proposta pela empresa Ziva Tecnologia e Soluções, além do RRPP, possuía o cálculo de rotas entre campi através de OSPF, ou seja, caso uma fibra do anel fosse rompida, seria necessário que o OSPF recalculasse as rotas entre campi, evento que leva um tempo na ordem de segundos.

Outra inconsistência encontrada no projeto original era que, ao utilizar simultaneamente RRPP e OSPF no anel central, a política do RRPP de bloqueio de interfaces faria com que o OSPF sempre reconhecesse este enlace como indisponível.

Desta forma, pode-se concluir que a proposta feita se enquadra nas exigências expostas pela STI, além de também proporcionar a implementação de uma nova arquitetura para o monitoramento da rede que reflete diretamente na confiabilidade e no desempenho da mesma.

O processo de implementação desta proposta se encontra em curso, estando com o anel integralmente renovado e com parte de seus capilares já adequados ao novo cenário.

A estrutura de monitoramento está sendo atualizada com o sistema de alarmes já em funcionamento e as outras plataformas em regime de testes.

7.1 Trabalhos Futuros e Continuidade

7.1.1 OSPF

Os experimentos realizados nesse trabalho abordaram apenas a camada de enlace. Também é preciso estudar a camada de rede, para tal camada, é proposto o OSPF. Em um breve estudo sobre o mesmo, pensou-se em separar as áreas onde o núcleo da rede seria a área zero e os capilares em conjunto com um *switch* do núcleo da rede configurariam as outras áreas.

Embora tenham sido feitos alguns estudos iniciais, ainda se faz necessário um estudo mais aprofundado sobre o OSPF, analisando, dentre outras características, o funcionamento de rotas redundantes e o tempo de convergência. Dessa forma este tema é válido para um futuro trabalho sobre o OSPF e sua implementação na rede da UFF.

7.1.2 Agregação de *Link*

Como visto na Figura 5.2 , faz parte da arquitetura proposta a agregação de *link*. Então é importante verificar como os clientes abaixo dos capilares reagem caso uma ligação a um dos *switches* em *stack* seja perdida.

Este trabalho não levou em consideração o tempo que a agregação de *link* leva para saber que um *link* está fora e que agora ele não deve enviar mais tráfego pelo mesmo. Essa situação pode acontecer por um rompimento de fibra ou caso um *switch* em *stack* seja desligado.

Existem algumas formas diferentes de como configurar essa redundância, em um estudo preliminar se optou por agregação de *link*, mas poderia ter sido escolhido uma fibra de backup apenas. Como esse estudo deve ser algo mais aprofundado a STI optou por aceitar a proposta de empresa. Isso não quer dizer que não seja necessário um estudo sobre o mesmo. Por esse motivo seria um ótimo tema para estudo.

7.1.3 *Quality Of Service* (QoS)

A rede da UFF necessita de QoS, pois todos os tráfegos de usuários acessando a internet está junto do tráfego de aplicações em tempo real, como *Voice Over IP* (VOIP) e videoconferência. Esses serviços são providos pela STI para integrar os cursos ministrados a distância e apresentações de teses com a participação de orientadores de outras regiões do país. Portanto se faz necessário que eles tenham uma boa qualidade.

Logo estudar como implementar QoS na atual rede da UFF é um trabalho importante que pode ser feito por um aluno de graduação em conjunto com a STI.

Bibliografia

[1] Huawei Technologies Co. *RRPP Technology White Paper*. Rel. téc. Out. de 2008.

[2] *Institute of Electrical and Electronics Engineers*. Disponível em `http://www.ieee.org`.

[3] ITU-T. *ITU-T Rec. G.8032/Y.1344 (02/2012) Ethernet ring protection switching*. Rel. téc. Disponível em `http://www.itu.int/rec/T-REC-G.8032-201202-I/en`. Out. de 2012.

[4] Sandie. *Cisco Resilient Ethernet Protocol White Paper*. Rel. téc. Set. de 2007.

[5] IEEE. *Rapid Spanning Tree Protocol*. 2004. URL: `http://standards.ieee.org/getieee802/download/802.1D-2004.pdf` (acesso em 07/01/2014).

[6] IEEE. *Multiple Spanning Tree Protocol*. 2004. URL: `http://standards.ieee.org/getieee802/download/802.1Q-2005.pdf` (acesso em 07/01/2014).

[7] Skype. *Skype*. 2003. URL: `http://www.skype.com/` (acesso em 06/01/2014).

[8] Alexey Kuznetsov. *Ping*. URL: `http://www.skbuff.net/iputils/` (acesso em 06/01/2014).

[9] NLANR/DAS. *Iperf*. URL: `http://iperf.sourceforge.net/` (acesso em 06/01/2014).

[10] tcpdump.org. *Tcpdump*. 2000. URL: `http://www.tcpdump.org/` (acesso em 06/01/2014).

[11] Gerald Combs. *Wireshark*. 1998. URL: `http://www.wireshark.org/` (acesso em 06/01/2014).

[12] Rob Hartill David Robinson Cliff Skolnick Randy Terbush Robert S. Thau Andrew Wilson Brian Behlendorf Roy T. Fielding. *Apache HTTP Server*. 1995. URL: `http://httpd.apache.org/` (acesso em 06/01/2014).

[13] Google Inc. *Chrome*. 2008. URL: `https://www.google.com/intl/pt-BR/chrome/browser/` (acesso em 06/01/2014).

[14] Ivan Pepelnjak. *udpflood.pl*. 2008. URL: `http://blog.ipspace.net/2008/03/udp-flood-in-perl.html` (acesso em 06/01/2014).