Helton Luiz Porto

Redundância e Balanceamento de Carga em Rede Corporativa

São José – SC
Março / 2014

Helton Luiz Porto

Redundância e Balanceamento de Carga em Rede Corporativa

Monografia apresentada à Coordenação do Curso Superior de Tecnologia em Sistemas de Telecomunicações do Instituto Federal de Santa Catarina para obtenção do diploma de Tecnólogo em Sistemas de Telecomunicações.

Orientador:
Prof. Eraldo Silveira e Silva, Dr.

CURSO SUPERIOR DE TECNOLOGIA EM SISTEMAS DE TELECOMUNICAÇÕES
INSTITUTO FEDERAL DE SANTA CATARINA

São José – SC
Março / 2014

Monografia sob o título "*Redundância e Balanceamento de Carga em Rede Corporativa*", defendida por Helton Luiz Porto e aprovado em 27 de fevereiro de 2014, em São José, Santa Catarina, pela banca examinadora assim constituída:

_____

Prof. Eraldo Silveira e Silva, Dr.

Orientador

_____

Prof. Marcelo Maia Sobral, Dr.

IFSC

_____

Karin Eickhoff Cavalhieri, Eng.

Gerência O&M Dados/Banda Larga - OI

Sempre que te perguntarem se podes fazer um trabalho,
respondas que sim e te ponhas em seguida a aprender como se faz.
F. Roosevelt

Agradecimentos

Dedico meus sinceros agradecimentos àqueles que muito me ajudaram para concluir este trabalho. Em especial minha esposa Maria Cecília Cardoso Porto por sempre me apoiar durante todo meu percurso acadêmico e profissional e meu orientador Dr. Eraldo Silveira em sua grande dedicação ao trabalho.

Dedico também aos meus colegas de faculdade, companheiros de trabalho, grandes professores, que com certeza me ajudaram de alguma forma e tornaram a realização deste trabalho em um sonho pessoal.

.

Resumo

Redundância e balanceamento de carga são características essenciais na maioria das redes corporativas. O uso de múltiplos acessos em uma rede "*stub*" é uma abordagem natural para construir redes robustas embora não garanta todos os tipos de falhas e não necessariamente permita fazer um balanceamento de carga adequado. Em geral, não é possível alcançar os dois recursos com um único protocolo de rede.

Este trabalho visa explorar o uso combinado desses dois sistemas utilizando de protocolos e recursos disponíveis em equipamentos de redes, a fim de obter um sistema que tolere falhas e aproveite todo o recurso disponível. Neste trabalho foi avaliado a integração de um protocolo de redundância de primeiro hop com o protocolo de roteamento BGP. Foram comparados três soluções para o protocolo de redundância de primeiro hop: HSRP, VRRP e GLPB. Com base nos resultados desta avaliação, foi escolhido o HSRP para ser integrado com o protocolo de roteamento BGP. O HSRP demonstrou um bom tempo de convergência quando falhas são injetadas na interface com a rede local e de acessos externos. O BGP foi utilizado para se conseguir o recurso de balanceamento de carga de entrada e de saída de tráfego da rede. Todos os experimentos foram realizados utilizando roteadores reais, porém em um ambiente experimental controlado.

Abstract

Fault tolerance and load balancing are essential features of most corporative networks. Multihoming stub networks are a natural approach to make a robust network but not enough to tolerate all spectrum of faults and to deal with load balancing. In general, it is not possible achieve both features with a single protocol. In this work we evaluate the integration of a protocol of first hop redundancy with the classical BGP protocol. First, we compare three solutions for next hop redundancy: HSRP, VRRP and GLPB. Based on the results of this evaluation we have choose HSRP to integrate with BGP protocol. HSRP has demonstrated good convergence time when faults are injected in the interface to the local network. We have used BGP to load balancing ingoing and outgoing traffic based on fixed traffic classes. All the experiments were carried out using real routers but in a controlled experimental environment.

Sumário

Lista de Figuras

Lista de Tabelas

Lista de Gráficos

1 Introdução

1.1 Motivação

Atualmente o mundo vem se modernizando, criando ferramentas para facilitar o cotidiano, possibilitando ganhar mais tempo e qualidade de vida. O avanço da tecnologia é uma das grandes responsáveis por essas novas conquistas e a cada dia nos tornamos mais dependentes destes serviços, principalmente na área de telecomunicações. O acesso à Internet e outros meios de comunicação estão sendo vistos como serviços triviais para a sociedade, principalmente no âmbito corporativo. Com esta dependência, Provedores de Serviços (ISP) ou operadoras que proveem serviços de conexão de acesso à Internet e comunicação de dados a seus clientes, vem se aprimorando em novas tecnologias, estudos e serviços para uma total disponibilidade.

No ambiente de redes de computadores, uma única conexão a uma rede externa (ISP), com um único caminho de saída, é comumente chamado de rede *stub*, e para que uma rede *stub* possua a capacidade de tolerar falhas na sua conexão lógica/física ao seu provedor (ou provedores), é necessário redundância de enlaces físicos. Essas redes que provêm de mais de um acesso a ISP, podem ser chamadas de redes *stub multi-homing*. Existem várias opções de redundância física que podem ser implementadas. De forma geral, pode-se ter (i) um *gateway* da rede *stub* conectado ao ISP via dois ou mais enlaces, ou ainda, (ii) dois ou mais *gateways* da rede *stub*, cada um com enlace para o ISP (ou diferentes ISPs).

Dentro das possibilidades acima pode-se ter um enlace primário (ativo) com um ou mais enlaces em estado de redundância passiva (*hot-standby*), ou ainda, pode-se ter todos enlaces ativos implementando-se alguma forma de balanceamento de carga. Em adição, o cliente proprietário da rede *stub* pode desejar aplicar diferentes estratégias para o tráfego que sai de sua rede (*outgoing traffic*) e para o que entra (*ingoing traffic*). Existem várias tecnologias e diferentes formas de implementar a redundância passiva e o balanceamento de carga. Por exemplo, é possível usar os protocolos que permitem redundância com a técnica de

duplicação de *gateway*, como o VRRP (*Virtual Router Redundancy Protocol*), um protocolo não proprietário descrito na RFC 3768, e os protocolos proprietários da Cisco, HSRP (*Hot Standby Router Protocol*) e GLBP (*Gateway Load Balancing Protocol*), que também provê balanceamento de carga no primeiro salto. Protocolos de roteamento também podem ser usados em conjunto com os protocolos mencionados como o OSPF e BGP, que também possuem mecanismos para implementar tais características de redundância e balanceamento, porém em camada de rede.

Por vivenciar frequentes situações de falhas no meu ambiente de trabalho, afetando as funções e o desempenho da empresa, minha proposta neste trabalho é criar um projeto de redundância para uma empresa corporativa. O projeto busca soluções de falhas de acesso a uma rede externa (*wan*), como por exemplo, acesso a internet e sistemas privados, tendo o aproveitamento total do sistema principal e redundante.

1.2 Objetivos

O objetivo geral do trabalho é avaliar o uso combinado de tecnologias existentes para redundância no primeiro *hop* (protocolos *HSRP, VRRP e GLBP)* com o protocolo de roteamento dinâmico BGP, em um cenário real de uma rede corporativa. As propriedades almejadas do sistema incluem além da tolerância a falhas o balanceamento de carga com fins de aproveitas os recursos de enlaces de reserva.

Em particular, pretende-se comparar, avaliar desempenho e levantar as limitações de cada combinação de protocolos bem como gerar recomendações para ajuste de parâmetros em determinados cenários específicos de uso.

1.3 Organização do Trabalho

O trabalho está organizado em mais quatro capítulos. No capítulo seguinte é descrita a fundamentação teórica do trabalho, explorando os protocolos que serão utilizados nos experimentos.

O capítulo 3 explora os protocolos de primeiro *hop* com experimentos práticos, a fim de obter suas avaliações. No capítulo 4, é montado um cenário ideal para o desenvolvimento do projeto almejado para uma empresa corporativa, fazendo experimentos e explorando sistema de redundância e balanceamento de carga, e também apresentando suas avaliações e conclusão final.

2 Fundamentação Teórica

Neste capítulo são apresentados conceitos básicos associados a tolerância a falhas e balanceamento de cargas em redes. Na sequência são apresentadas algumas soluções tecnológicas para a implementação de redundâncias e do balanceamento de cargas em nível de camada física/enlace e de camada de rede. Inicialmente são exploradas soluções para redundância para o roteamento no primeiro salto. Posteriormente são apresentadas soluções para implementação de redundância e balanceamento em cenários de múltiplos enlaces de uma rede corporativa. O uso do BGP é destacado neste ponto devido ao interesse deste trabalho de integrá-lo com protocolos de redundância de primeiro salto.

2.1 Alguns conceitos associados à tolerância a falhas e balanceamento de carga

No ambiente de redes de computadores, sempre se estará sujeito a falhas do sistema, indisponibilizando os recursos oferecidos que muitas vezes são de extrema importância para seus usuários. Neste sentido, a concepção de um sistema de alta disponibilidade envolve o uso de técnicas que se aplicam a prevenção de falhas, a tolerância a falhas, a remoção de falhas e a predição de faltas (AVIZIENIS, 2004).

A **tolerância a falhas** diz respeito à propriedade de um sistema continuar a fornecer um serviço correto para o qual foi projetado, mesmo quando submetido à faltas de *hardware* e *software*. Esta propriedade pode ser obtida pela aplicação de técnicas de redundância específicas, tais como de *hardware* e de *software*.

Redundância

A redundância é definida como a "capacidade de um sistema em superar a falha de um de seus componentes através do uso de recursos redundantes" (PINHEIRO, 2004). Para isso, um sistema depende de recursos alternativos, além do principal, e que estejam disponíveis para assumir o sistema assim que um evento de falha ocorrer.

A redundância é um termo muito usual e não se aplica apenas a redes de computadores.

Ela pode ser embarcada em vários sistemas, como: energia, aviação, máquinas industriais e outros. A finalidade de um recurso redundante é suprir integral ou parcialmente os serviços, visando sempre que as funções mínimas do sistema estejam em funcionamento.

Contingência

PINHEIRO (2004) define contingência como "possibilidade de um acontecimento futuro de uma condição existente, incerteza sobre as condições operacionais envolvidas e a resolução destas condições dependerem de eventos futuros". Ou seja, a possibilidade de um fato ocorrer ou não, com uma situação de risco existente, com certo grau de probabilidade de acontecer.

Portanto, quando se projeta um sistema, deve-se definir em conjunto um plano de contingência, plano este que deve visar manter o funcionamento do sistema, caso ocorra uma falha, quando este estiver executando tal processo de contingenciamento.

Para isso, faz-se necessário um estudo de cada um dos processos em particular, quais os riscos envolvidos em cada um deles, de como afetariam o sistema, quais seriam os mais impactantes, quais as áreas mais críticas, o que poderia paralisar o sistema, e o tempo de restabelecimento para cada fase, pois são questões que norteiam o plano de contingência.

Medidas preventivas e planejadas que suportem, por exemplo, falhas de *software*, *hardware*, base de dados, energia, temperatura, perda do *link* de comunicação e de causas naturais, devem estar incluídas no plano de contingenciamento, ou seja, ações imediatas, para serem executadas, visando o restabelecimento dos serviços, mesmo que parcialmente, diminuindo o tempo de paralisação caso ocorra uma falha.

O plano de contingência deve ter alta disponibilidade de informações de monitoramento. Ser implantado de um modo seguro e eficiente, que possa gerenciar/solucionar os problemas ocorridos, e se possível, ser pró-ativo e disponibilizem a solução da falha independentemente de ações externas, minimizando os impactos, e apenas mantendo relatório dos fatos ocorridos.

Disponibilidade

Disponibilidade é definida pelo tempo em que um sistema de rede deve estar disponível para seus usuários (PINHEIRO, 2004). Ela pode ser mensurada em relação ao tempo em que o sistema está em falha (*downtime*), com o tempo que deve estar disponível. Dependendo do plano de contingência criado para suprir falhas que possam indisponibilizar o sistema, o tempo disponível pode variar em horas, dias, meses ou até anos.

A figura 1 representa uma tabela do tempo de falha de um sistema, em relação a um ano de operação da uma rede. Uma pequena variação na porcentagem pode considerar uma grande diferença de tempo. Por isso, é importante estimar a disponibilidade mínima da rede, a fim de montar seu plano de contingência. Segundo PINHEIRO (2004), a disponibilidade pode ser enquadrada em três classes, Disponibilidade Básica, Alta Disponibilidade e Disponibilidade Contínua.

Availability	Downtime per Year (24x7x365)		
99.000%	3 Days	15 Hours	36 Minutes
99.500%	1 Day	19 Hours	48 Minutes
99.900%		8 Hours	46 Minutes
99.950%		4 Hours	23 Minutes
99.990%			53 Minutes
99.999%			5 Minutes
99.9999%			30 Seconds

Figura 1: Exemplo de Disponibilidade

Balanceamento de Carga

Por ser um dos pontos mais críticos e vulneráveis a falha, a multiplicação de *links wan* é muito comum em planos de contingência. Dependendo do projeto e a disponibilidade dos recursos dos equipamentos utilizados, o balanceamento de carga entre os *links* pode ser implementado. Com o balanceamento de carga, pode-se aproveitar os recursos do sistema redundante, ao invés de ficarem ociosos até que ocorra uma falha.

A função de balanceamento entre os *links wan* é distribuir o tráfego de dados entre eles. Dependendo de sua aplicação, o balanceamento aumenta o desempenho da rede somando a banda dos *links*, aumentando a capacidade do sistema podendo inclusive prover redundância entre os links. O balanceamento pode ser relativo ao tráfego que entra na rede e ao tráfego que sai, ela pode ser em nível de pacotes, fluxos, destinos, e entre outras possibilidades.

2.2 Tecnologias para Redundância de acesso no primeiro *hop*

A **multiplicação de *gateway*** padrão é umas das tecnologias mais usadas para redundância no primeiro *hop*. Este recurso é utilizado através de protocolos de rede, e fazem com que a rede fique protegida por dois ou mais equipamentos *gateway* da rede (CPE) caso ocorra uma falha no equipamento e também nas interfaces de acesso a rede externa. Deve ser observado que em geral os hospedeiros da rede não possuem mecanismos de descoberta de

novo *gateway*. Por exemplo, em configurações usando o serviço DHCP, os hospedeiros recebem o endereço IP juntamente com informações adicionais tais como servidor de DNS e *gateway default*. Em uma rede local tal como a *ethernet*, a descoberta do endereço MAC do *gateway default* se dá pelo uso do protocolo ARP. O hospedeiro solicita por meio de *broadcasting ethernet* que o detentor do endereço IP responda e desta forma ela descobre seu endereço MAC. Como será visto adiante, a virtualização de endereços MAC e endereços IP é a base da construção de protocolos de redundância de primeiro salto.

Os roteadores integrantes de um grupo específico trocam mensagens entre eles com várias informações, que podem transferir o estado do *gateway* em espera como ativo ou vice-versa. Os protocolos estudados são o HSRP, VRRP e o GLBP.

O protocolo HSRP

O HSRP (*Hot Standby Router Protocol*) é um protocolo de rede proprietário definido pela RFC 2281 desenvolvido pela Cisco. Aplicado em um domínio de rede, em dois ou mais roteadores, cada roteador terá seu ip fixo da rede e também o ip *gateway* padrão do domínio, comum em todos os roteadores, e chamado de ip virtual. Desta forma, o *gateway* dos dispositivos finais da rede não precisa ser alterado. Os roteadores trocam mensagens *hello* a cada 3 segundos através do endereço *multicast* 224.0.0.2, usando a porta UDP 1985. Para definir o *gateway* ativo (*link* principal) ou *stand-by (link back-up)* no domínio, são definidos prioridades nos roteadores. O roteador que tiver maior prioridade é eleito como ativo. O valor padrão da prioridade é 100, e se não definido, o roteador que tiver o maior valor ip é eleito ativo, mas geralmente se define a prioridade estaticamente por escolha do administrador da rede.

O HSRP é configurado nos roteadores para que se houver uma falha em alguma regra estabelecida no roteador, como queda de interface ou falta de conectividade no próximo *hop* (WAN), decrementa-se o valor de sua prioridade, a fim do roteador vizinho se tornar ativo. Caso normalize as regras, o roteador dado como *link* principal, volta a sua prioridade inicial e seu estado como ativo.

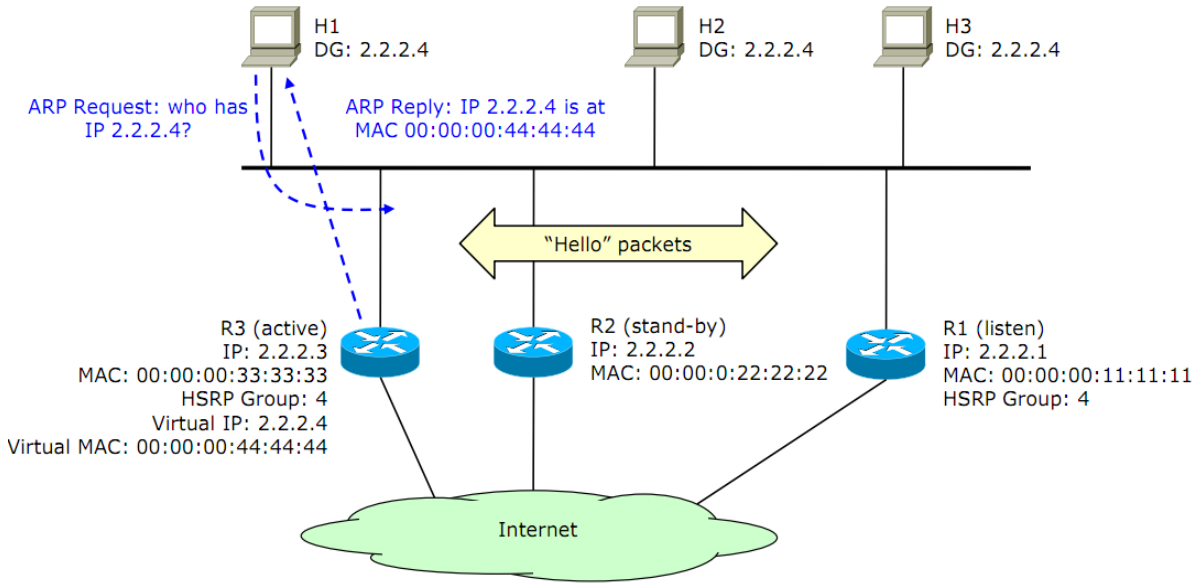

Figura 2: Exemplo de Funcionamento do HSRP (RISSO, 2013)

A Figura 2 ilustra o funcionamento do HSRP. Note-se que os hospedeiros estão configurados para um *gateway default* 2.2.2.4. Este endereço é virtual. O roteador R3, integrante de um grupo de roteadores HSRP e na condição de roteador ativo, responde por requisições ARP. O endereço de MAC fornecido é virtual (00:00:00:44:44:44). Em caso de falha de R3, o roteador R2, até então em *stand-by* , assume a condição de ativo e o roteador R1 passa a ser *stand-by*. R2 passa então a responder pelo IP virtual e pelo MAC virtual.

As seguintes considerações ainda podem ser realizadas sobre o HSRP:

- cada VLAN é uma LAN separada e portanto utiliza-se cada uma de um *gateway default*. Neste caso são requeridos múltiplos grupos HSRP;
- apenas o roteador ativo se utiliza do MAC virtual nos pacotes de *HELLO*. Desta forma, quando muda de roteador ativo, se existirem *switches* no caminho, eles aprenderão onde está o novo roteador. Neste sentido, o novo roteador ativo também emite um ARP *reply* em *broadcast*.
- apenas o enlace ligado de saída do roteador ativo é usado. Os enlaces dos demais roteadores ficam desperdiçados. Para tráfego que entra, todos os *links* podem ser utilizados, permitindo desta forma assimetria no tráfego. Note-se que o HSRP **não influencia** neste processo, cabendo se for o caso, a um protocolo de roteamento tal como o BGP.

Finalmente, é importante ressaltar que o HSRP não possui mecanismos para balanceamento de tráfego que sai. É possível, no entanto, criar grupos HSRP diferentes, cada um com um *gateway* ativo definido, de forma que hospedeiros com IP virtual de um grupo encaminham por um *gateway* enquanto outros hospedeiros encaminham por outro *gateway default*.

O protocolo VRRP

O protocolo VRRP é muito semelhante ao HSRP, porém é um protocolo aberto e definido pela RFC 3678, podendo ser usado em qualquer equipamento que suporte o protocolo. O VRRP diferente do HSRP define o status dos roteadores como *master* para o ativo e *backup* para os demais utilizando também do valor de prioridades em cada equipeamento. As mensagens trocadas pelos roteadores são chamados de *Link-State Advertisement* (LSA), através do endereço *multicast* 224.0.0.18. Caso ocorra alguma falha no roteador mestre, o mesmo envia mensagens para os roteadores *backup* e um é eleito a assumir o status de *master*.

O protocolo GLBP

O GLBP (*Gateway Load Balancing Protocol*), também é um protocolo proprietário da Cisco, semelhante aos protocolos já mencionados, porém além de fazer redundância de rede, ele provê balanceamento de carga. A principal diferença do GLBP em relação aos demais é que ele consegue atribuir diferentes endereços MAC para um mesmo IP Virtual (KRAEMER, et al., 2010, p. 2). Os roteadores eleitos back-up também são encaminhadores e ativos.

No GLBP existem dois tipos de *gateways* ativos: o *Gateway* Virtual Ativo (AVG) e o *Gateway* Virtual Encaminhador (AVF). O AVG é eleito pelo grupo e os AVF são seus *backups*. A cada solicitação ARP feita ao AVG é devolvido o MAC Virtual de outro roteador AVF. Com este mecanismo, o endereço MAC do *gateway* armazenado na tabela ARP do cliente não é o mesmo em todas as estações, permitindo o balanceando da carga [Satapati et al.,2004, pp.02].

2.3 Redundância e Balanceamento com foco em Múltiplos Enlaces *Wan*

Existem vários cenários, diferentes topologias e meios de acesso à rede. A multiplicação de *links wan* é mais comum para a aplicação num plano de contingência. Ela pode ser fornecida por um único ou diferentes provedores de acesso. Múltiplas conexões são conhecidas como redes *multi-homing* e aplicadas de várias maneiras. Iremos explorar duas

topologias diferentes *multi-homing*.

***Multi-homing* com CPE único**

Esta configuração permite ter mais de um acesso *wan*, porém os *links* são concentrados em apenas um CPE. A redundância fica apenas no enlace entre o cliente e operadora. Nesta topologia os protocolos de duplicidade de *gateway* (HSRP, VRRP e GLBP) não são aplicáveis, porém se utiliza dos protocolos de enlace e roteamento para prover redundância. Nos acessos simétricos, em que os *links* são iguais, pode ser aplicado balanceamento de carga entre os circuitos. Em acessos assimétricos, ou seja, *links* de diferentes velocidades, é mais comum que se utilize o plano de *links* principal(ativo) e *backup*(ocioso).

***Multilink* PPP**

O *Multilink PPP* (MP)é um protocolo da camada de enlace descrito na RFC 1990. Ele tem função de agregar dois ou mais circuitos seriais que utilizam o protocolo PPP. Com a aplicação deste protocolo, fica o CPE com apenas uma interface lógica de saída, porém agregado a vários circuitos físicos. A vantagem do MP é que ele soma a banda de cada circuito, ou seja, se o CPE tiver disponível 3 enlaces de 2Mbps, a banda total de saída será de 6Mbps. Outra vantagem é que se um *link* fica inoperante não há indisponibilidade do sistema, apenas a banda diminui para a quantidade de *links* ativos. A desvantagem desta topologia é que não há redundância de CPE e PE, pois os *links* precisam estar em um único CPE e também na operadora (PE), os *links* também devem ter a mesma velocidade.

Balanceamento por pacotes

O balanceamento por pacotes é feita através da camada de rede. O tráfego é divido pelas interfaces de saída do CPE, balanceando pacote por pacote em cada enlace, distribuindo uniformemente o tráfego de saída. A tabela de roteamento do CPE é montada com duplicidade de saída para o mesmo destino. Nesta aplicação também é somada a banda de cada circuito. Algumas desvantagens são relevantes nesta topologia. O grande uso de CPU no CPE, podem ocorrer problemas com aplicações que não toleram atrasos e perdas e os *links* devem ser simétricos. Esta configuração também exige uma configuração complexa para total disponibilidade caso ocorra falha em algum circuito. Como o roteamento é estático e não há verificação do estado de enlace, ou seja, o roteador pode distribuir o tráfego em um enlace que está em falha, causando perdas ou total indisponibilidade na rede.

Balanceamento por destino

O balanceamento por destino é basicamente a configuração default dos equipamentos de rede (CPE e PE). Para um funcionamento correto, os circuitos também devem ser simétricos, e a diferença básica para o balanceamento por pacote, é que cada fluxo de dados é encaminhado por um único *link*, ou seja, a distribuição não é feita por pacote, e sim, por controle de fluxo TCP. Por exemplo, um *download* feito na rede com este recurso, o tráfego entra por apenas um *link*, um outro *download* simultâneo entraria por outro *link*, e assim sucessivamente. A vantagem desta configuração é o baixo uso do CPU dos roteadores e o bom funcionamento para tráfegos multimídia UDP. A grande desvantagem é que não é somada as velocidades de cada circuito. Se o sistema, por exemplo, fornece 3 *links* de 2Mbps, o usuário conseguirá fazer três *downloads* com banda de 2Mbps cada, e não um de 6 como nas outras topologias.

Acessos Diferenciados

É quando se utiliza de diferentes tecnologias de acesso para redundância. Este sistema geralmente se utiliza de um *link* principal e outro de back-up, que fica ociosa na espera do *link* principal falhar a assumir o tráfego. Os *links* podem ser simétricos com o objetivo de não alterar o desempenho da rede, ou inferior, apenas para suprir as necessidades básicas da rede utilizando um acesso e menor custo. Um exemplo é o sistema ter um acesso principal com um circuito de dados dedicado de alto desempenho, e um *backup* com tecnologias frame-relay, adsl ou até um acesso discado. A comutação do roteamento é feita no CPE, criando regras através do status de interface ou falta de conectividade do *link*. A vantagem é de não depender de uma única tecnologia de acesso e também de diferentes operadoras. Esta topologia é uma das mais usadas em planos de contingência devido ao seu menor custo de implementação.

Multi-homing com múltiplos CPE's

Este é um tipo mais completo de redundância que pode ser aplicado. Além de múltiplos acessos *wan* para operadoras diferentes, há também a redundância de CPE conforme ilustrado na figura 3. Nesta, é aplicado os recursos de multiplicação de *gateway*, através dos protocolos HRSP, VRRP e GLBP mencionados no trabalho.

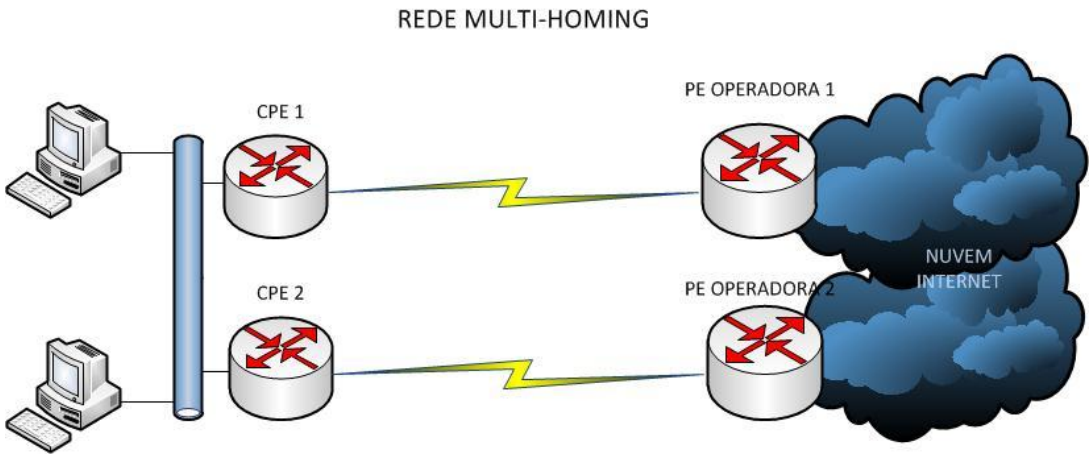

Figura 3. Rede *Multi-homing*

2.4 O protocolo BGP e a aplicação de políticas no uso de enlaces

O protocolo BGP

O BGP (*Border Gateway Protocol*) é um protocolo de roteamento dinâmico utilizado por operadoras para interconexão entre sistemas autônomos (*Autonomous Systems - AS*). Porém, ele é comumente utilizado em redes privadas para transportes de tabela de roteamento e por seus vários recursos. Em cada enlace há uma sessão bgp ativa com o roteador vizinho *(next-hop)*, enviando e recebendo as tabelas de roteamento dinamicamente.

O BGP é um protocolo utilizado para a troca de informações de roteamento entre sistemas autônomos da Internet. Atualmente é o único protocolo utilizado para este fim. A última versão do BGP é conhecida como BGP4 (RFC4271).

A troca de informações é realizada pelos roteadores de borda dos sistemas autônomos através de sessões TCP (porta 179). Uma nuvem BGP pode ser vista como um conjunto de supernós interligados por *links* virtuais. O protocolo BGP se aproxima da abordagem por vetor de distâncias. No entanto, ele fornece um caminho completo de sistemas autônomos (AS's) que compõe o caminho para uma determinada rede de destino informada. A métrica usada é o *hop* em nível de AS.

Roteadores que falam BGP (BGP *speakers*) podem estar conectados entre dois AS's diferentes (EBGP ou *external* BGP) ou podem conversar internamente a um AS (IBGP ou Internal BGP).

Atributos do BGP

O uso de alguns atributos das mensagens do BGP, combinados com o seu algoritmo de seleção de rotas, é um instrumento valioso de controle no uso de enlaces e na aplicação de políticas de uso da rede. Pode-se destacar os seguintes atributos:

Atributo AS-PATH: Quando o anúncio de uma rota é publicado por um sistema autônomo, o número de AS deste sistema é adicionado à lista de números AS que o anúncio possui. O AS-PATH será utilizado para escolha dos caminhos, isto é, se um AS recebe anúncios da mesma rede através de diferentes AS vizinhos, então ele pode escolher (dentre outros critérios) o anúncio com menor AS-PATH. É um comportamento similar ao algoritmo de vetor de distância.

Um AS-PATH também pode ser usado para dar preferência a entrada de tráfego por um determinado caminho do AS. O AS insere múltiplas cópias de seu AS-ID no ASPATH para "enganar" o AS a ele conectado. Abaixo, a figura 4 ilustra o uso de pre-pending no AS-PATH.

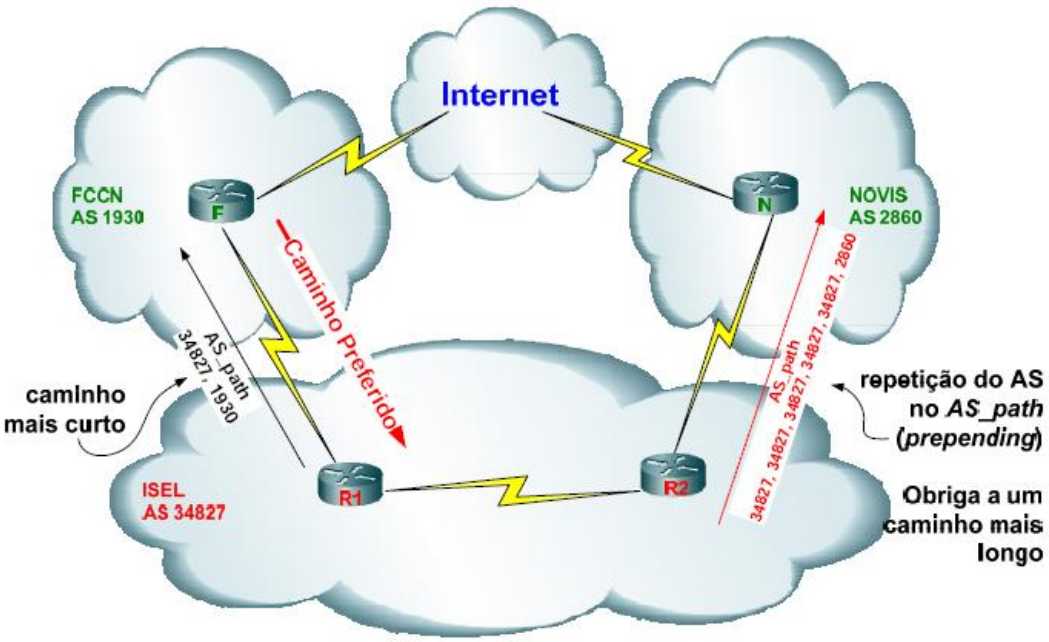

Figura 4: Uso de Pre-Pending no AS-PATH. Fonte: (ISEL, 2013)

Atributo *Local Preference*: Este atributo é utilizado para dar preferência a um caminho de saída do sistema autônomo. Ele é propagado dentro do mesmo AS ilustrado na figura 5.

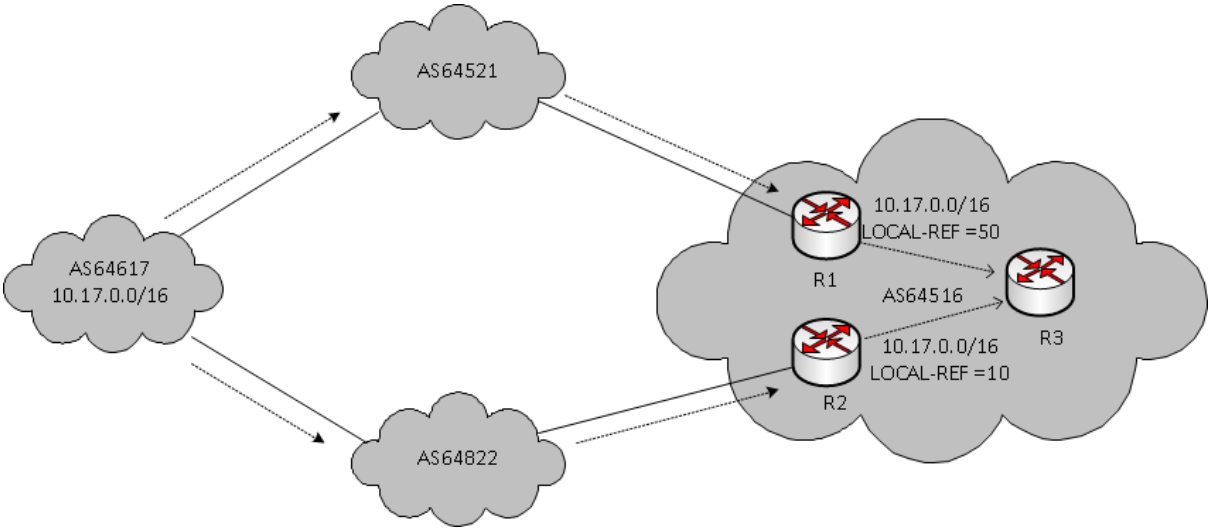

Figura 5: Uso do Atributo Local-Preference.

Atributo MED: este atributo é enviado como sugestão a um AS externo para dar preferência a um dos caminhos entre dois sistemas autônomos. Trata-se somente de uma sugestão porque é o AS externo que decide levando em consideração outros atributos. A figura 6 ilustra o uso do parâmetro MED.

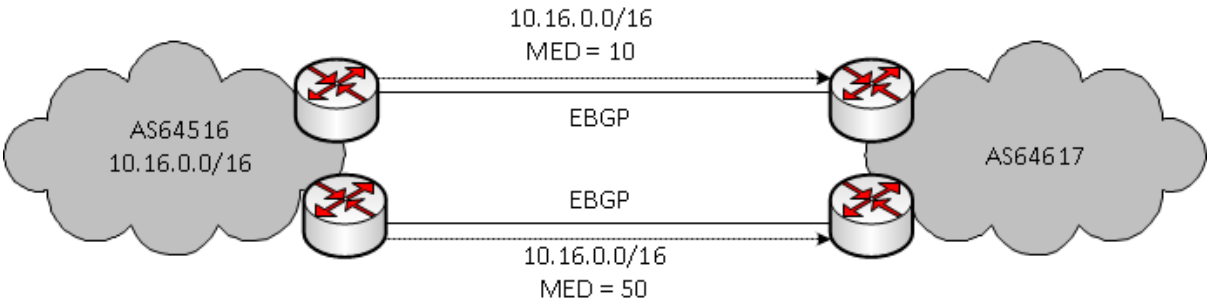

Figura 6: Uso do Parâmetro MED.

Atributo peso: Se um *router* aprende mais do que uma rota para o mesmo destino, a rota com o maior peso é utilizada como mostra a figura 7. É um atributo proprietário da Cisco. Este atributo não é anunciado a outros roteadores.

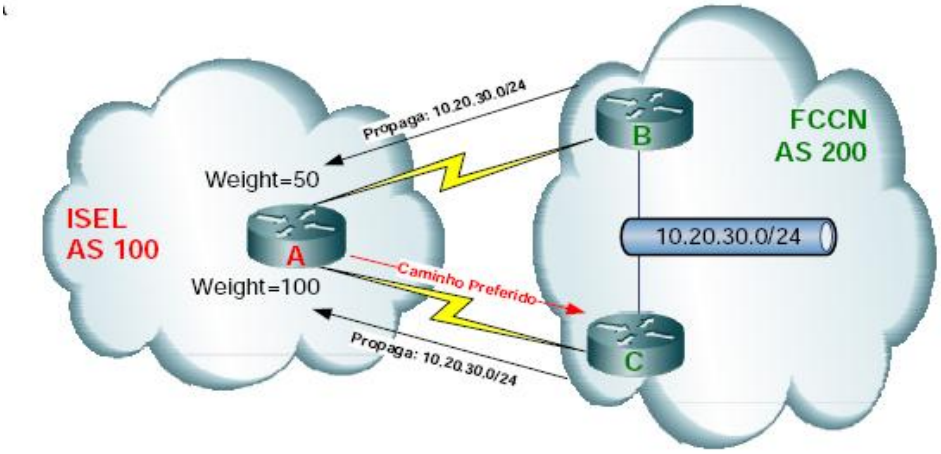

Figura 7: Uso do Parâmetro Peso. Fonte: (ISEL, 2013)

Políticas de Uso

No BGP são criadas regras para que as tabelas de roteamento fiquem duplicadas para cada enlace ou sessão BGP. Se não for aplicada nenhuma métrica ou custo na tabela, o tráfego pode ser balanceado entre os *links*, tanto por destino, quanto por pacotes. Além do balanceamento, o BGP permite administrar o tráfego de entrada e saída, possibilitando o tráfego sair por um *link* e entrar por outro, aplicando regras com custos, métricas, AS path prepending, local-prefence, mapeamento de rotas, tal como colocado anteriormente.

O BGP fornece também uma série de filtros que podem ser aplicados às redes para um determinado AS. Por exemplo, existem filtros por prefixo que permitem determinar que um grupo de prefixos de rede somente podem ser recebidos por determinados AS's. Existem também filtros aplicáveis conforme o atributo AS-PATH, permitindo selecionar caminhos que priorizem a passagem por determinados ASs.

São estas características que fazem do BGP um protocolo universalmente aceito para troca de informações de roteamento entre sistemas e que possibilitam até mesmo a sua aplicação em redes nas bordas do sistema.

2.5 Integração de Balanceamento de Carga nas estruturas de acesso

A integração de balanceamento entre *links* alternativos de acesso é muito explorado por operadoras e clientes, pois se permite um melhor aproveitamento dos recursos oferecidos. Um *link* ocioso, apenas aguardando um evento de falha ocorrer para enfim ser utilizado, pode ser considerado um desperdício de utilização de banda. O balanceamento entre os *links* pode

melhorar muito o desempenho de um sistema, já que nesta estrutura se soma a banda dos circuitos.

Utilização integrada de protocolos de multiplicação (VRRP, HSRP e GLBP) de *gateway* com BGP

Geralmente, os protocolos de multiplicação de *gateway* são configurados para que se utilize o estado das interfaces *wan* de camada 2 para leitura do evento da falha, decrementando sua prioridade, e enfim enviar as mensagens de status para os roteadores vizinhos assumirem o sistema. Um grande problema, pois nem sempre a falha pode estar ligada ao um problema de físico no acesso ou na camada 2, pois a conectividade pode ser perdida sem que o protocolo caia, por exemplo. A utilização do BGP pode ser aplicada na leitura da falha pelos protocolos através do status da sessão BGP, pois se não há conectividade com o próximo *hop*, a sessão BGP cai (*down*), ou seja, a parametrização da falha estará sendo feita acima na camada 3 de rede, utilizando os recursos de "*trackinkg*". O *track* é um recurso dos roteadores para rastreamento de um objeto ou evento, que permite o acompanhamento de determinados objetos específicos, tomando medidas quando o estado do objeto rastreado sofrer alteração, tais como queda da interface *wan*, perda de comunicação em um determinado destino, entre outros. Operadoras utilizam deste sistema para uma melhor performance nos serviços de redundância.

2.6 Conclusão

Neste capítulo foi mostrado todo embasamento teórico que servirá de referencia na execução do objetivo do trabalho. Foram mostrados alguns conceitos de redundância, contingência, tolerância a falhas e balanceamento de carga. Também foi explorado os protocolos de redundância de primeiro hop (HSRP, VRRP e GLBP) juntamente com protocolo de roteamento BGP, a fim de integrar os dois sistemas de redundacia e balanceamento de carga.

3 Experimento e Avaliação dos protocolos de redundância de primeiro Hop

Neste capítulo, serão explorados os protocolos de redundância de primeiro hop ou multiplicadores de gateway como são conhecidos. Será feito experimentos com equipamentos reais com os três protocolos citados neste trabalho HSRP, VRRP e GLBP. Os três protocolos basicamente exercem as mesmas funções, porem cada protocolo tem suas particularidades de funcionamento e configuração. Suas características podem impactar no desempenho de um sistema de redundância e será fundamental na escolha para sua aplicação.

3.1 Objetivo do Experimento

O objetivo do experimento é implantar um cenário utilizando os três protocolos com equipamentos reais em bancada, a fim comparar e avaliar o desempenho de cada protocolo. Espera-se obter o melhor desempenho dos protocolos fazendo o uso de ajustes nos parâmetros de configuração que possam alterar seu funcionamento padrão. O tempo de convergência será utilizado na comparação e avaliação dos protocolos e servirá como base para definir o protocolo mais rápido e eficaz dependendo de suas aplicações. Através dos resultados obtidos, será definido o protocolo utilizado no projeto de um cenário real, na simulação de uma rede corporativa.

3.2 Cenário e montagem do Experimento

A figura 8 ilustra o cenário utilizado nos experimentos. Foram utilizados equipamentos reais e montados em bancada. Os equipamentos foram: 2 roteadores Cisco 1700 (FNS1 e FNS2), um switch *layer* 2 Encore, cabos de conexão ethernet e 1 computador.

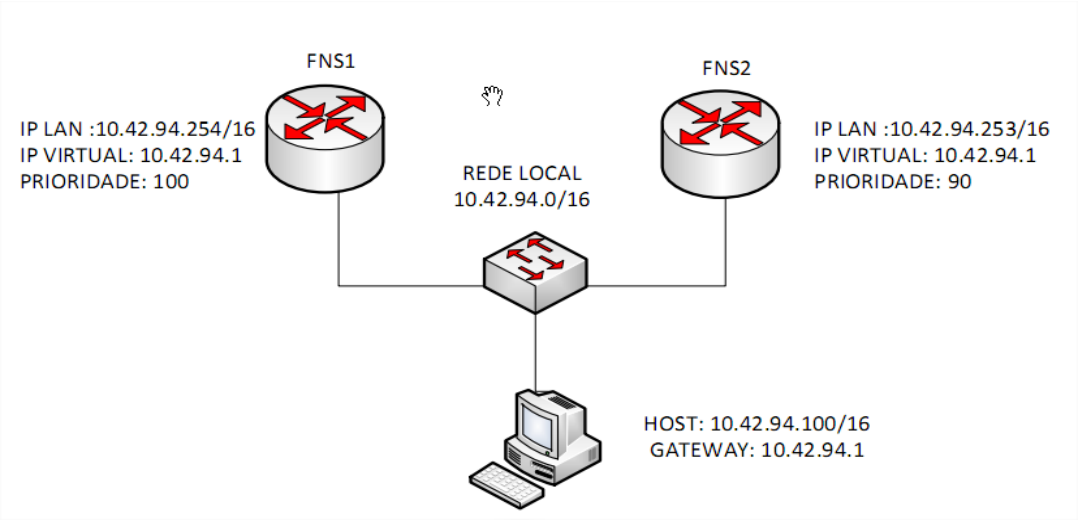

Figura 8: Cenário Primeiro *Hop*

Nos testes, o roteador FNS1 sempre será definido como o *gateway* principal da rede, e FNS2 back-up. A definição é feita através da configuração de prioridades nas interfaces *fastethernet* de cada roteador assim como o *gateway* da rede (10.42.94.1) conhecido como ip virtual, onde apenas o roteador ativo responderá por ele. Percebe-se que cada interface também recebe um ip ativo na interface, servindo de comunicação entre os dois roteadores nas mensagens dos protocolos (*hello/LSA*).

Abaixo seguem as configurações utilizadas no experimento para cada protocolo (Tabela 1). Percebe-se que o roteador principal FNS1 não apresenta configuração de prioridade, assim o roteador define a prioridade 100, padrão nos três protocolos. No roteador FNS2, foi definido prioridade 90, definindo o mesmo como o roteador back-up da rede, ficando apenas no status de leitura(*speak*). O comando "*preempt*" serve para definir se a falha no roteador for solucionada, o sistema retorne para estado original.

ROTEADOR FNS1	ROTEADOR FN2
Protocolo HSRP	**Protocolo HSRP**
interface FastEthernet0/0 description LAN_FNS1 ip address 10.42.94.254 255.255.0.0 no ip route-cache cef speed 100 full-duplex standby 1 ip 10.42.94.1 standby 1 preempt	interface FastEthernet0 description LAN_FNS2 ip address 10.42.94.253 255.255.0.0 speed auto full-duplex standby 1 ip 10.42.94.1 standby 1 priority 90 standby 1 preempt
Protocolo GLBP	**Protocolo GLBP**
interface FastEthernet0/0 description LAN_FNS1 ip address 10.42.94.254 255.255.0.0 speed 100 full-duplex glbp 1 ip 10.42.94.1 glbp 1 preempt	interface FastEthernet0 description LAN_FNS2 ip address 10.42.94.253 255.255.0.0 speed auto full-duplex glbp 1 ip 10.42.94.1 glbp 1 priority 90 glbp 1 preempt
Protocolo VRRP	**Protocolo VRRP**
interface FastEthernet0/0 description LAN_FNS1 ip address 10.42.94.254 255.255.0.0 speed 100 full-duplex vrrp 1 ip 10.42.94.1 vrrp 1 preempt	interface FastEthernet0 description LAN_FNS2 ip address 10.42.94.253 255.255.0.0 speed auto full-duplex vrrp 1 ip 10.42.94.1 vrrp 1 priority 90 vrrp 1 preempt

Tabela 1: Configuração

3.3 Execução

O experimento se baseia em simular uma falha no roteador principal e verificar o tempo de convergência de cada protocolo. Inicialmente pretendia-se utilizar como parâmetro na causa da falha um evento externo da rede local com o recurso de “tracking”. Porém, nas versões de software dos roteadores, o protocolo VRRP não suporta esta aplicação. Neste caso, o experimento se baseia em simular a queda física da interface lan do roteador principal (FNS1) com o switch, também simulando sua total indisponibilidade, cenário comum entre os três protocolos ilustrado na figura 9.

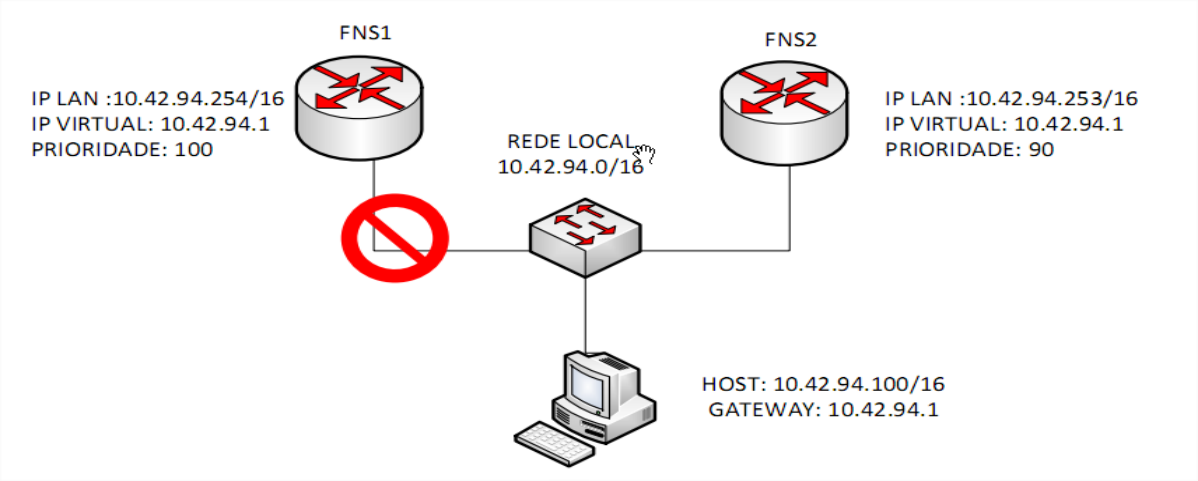

Figura 9: Falha no enlace LAN

Para análise do tempo de convergência de cada protocolo, foi sincronizado o relógio dos roteadores através do protocolo NTP e montado uma tabela de tempo através dos registros de logs dos roteadores em cada evento. No caso o roteador FNS1, registra o exato momento da queda de sua interface, consequentemente o roteador FNS2 registra o momento em que sua interface passa seu estado como ativo da rede. Com a tabela montada, é subtraído o tempo que roteador FNS2 assumiu o tráfego da rede com a queda da interface lan de FNS1, obtendo o tempo total do processo de convergência. A tabela completa é mostrada no anexo 1 e um exemplo de logs na figura 10 abaixo.

```
FNS1#
Oct  1 01:51:22.823: %LINK-5-CHANGED: Interface FastEthernet0, changed state to administratively down
Oct  1 01:51:23.823: %LINEPROTO-5-UPDOWN: Line protocol on Interface FastEthernet0, changed state to down


FNS2#
Oct  1 01:51:28.367: %HSRP-5-STATECHANGE: FastEthernet0 Grp 0 state Standby -> Active
```

Figura 10: Registro de Logs

Além do tempo, foi também analisado a quantidade de perdas de pacote para o gateway padrão da rede(10.42.94.1). Paralelamente o computador conectado a rede, simula um teste de conectividade com seu gateway, através de um *ping*. Foi utilizado o *ping* padrão do Windows de 32 bytes com o parâmetro "-t" tornando contínuo. Segue um exemplo na figura 11, onde nos mostra um pacote perdido no momento da falha da interface de FNS1, e seguindo os testes após o roteador FNS2 assumir a trefego.

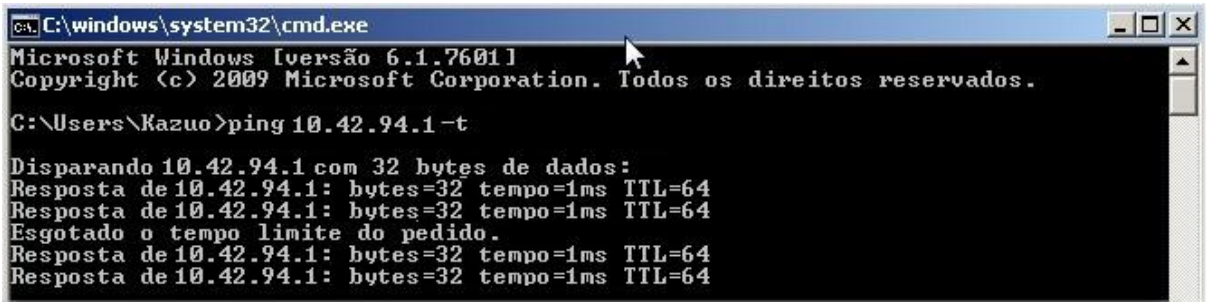

Figura 11: Ping

A figura 12 mostra os resultados obtidos através dos experimentos avaliados com a média de tempo de convergência e a média da quantidade de perdas de pacotes de cada protocolo. Foram executados 10 simulações de interrupção em cada protocolo, com intuito de perceber um padrão de comportamento.

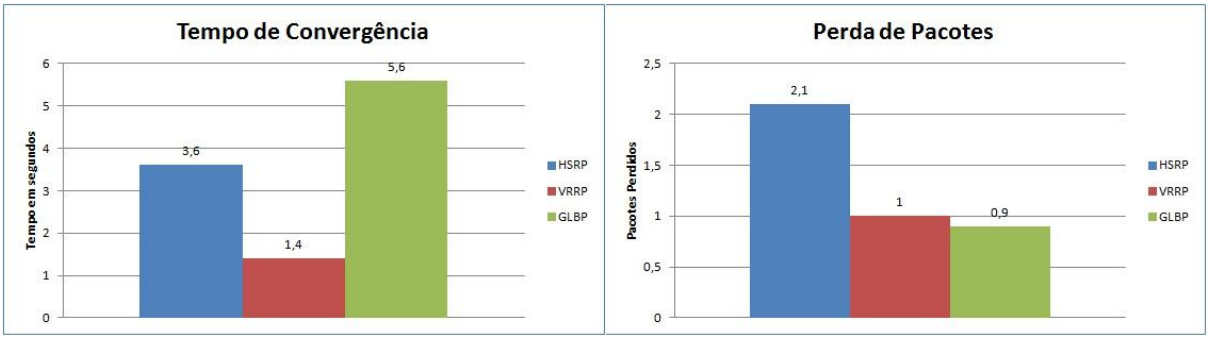

Figura 12: Média de tempo e pacotes perdidos

3.4 Avaliação dos Resultados

Através dos dados coletados no experimento, foi verificado que o protocolo VRRP comportou-se como o mais rápido dos 3 avaliados na média de 1,4 segundos de tempo de convergência, seguido do HSRP na média de 3,6 segundos e com o pior desempenho o protocolo GLBP. Os dados foram coletados com as configurações padrões de cada protocolo. A fim de experimento, foram feitos alguns ajustes nos tempos das mensagens nos protocolos (*hello*/SLA), tentando obter uma melhor performance. Porém, não obteve mudanças significativas nos seus desempenhos.

No desempenho de perda de pacotes com ping padrão para o *gateway* virtual da rede (10.42.94.1), o protocolo HSRP foi o que teve maior perda nos testes, na média de 2,1 pacotes perdidos, seguido do VRRP que teve uma média de 1 pacote perdido em cada simulação. O GLBP teve o melhor desempenho nos testes de conectividade, em alguns testes não houve perda. Isso se deve a uma característica do GLBP em fornecer múltiplos endereços MAC-ADDRESS, não havendo necessidade de atualização da tabela *arp* no processo da falha, o que não acontece com os protocolos VRRP e HSRP.

3.5 Conclusão

Os testes foram conclusivos para definir um protocolo de redundância de primeiro *hop*, na implementação de um projeto com a integração do protocolo BGP simulando um sistema de uma rede corporativa que será vista no próximo capítulo. O protocolo ideal seria o mais rápido nos testes e aberto como o VRRP, porém foi escolhido o HSRP. Um dos motivos é por não possuir balanceamento em camada 2 como no GLBP, e ser mais completo que o VRRP na questão de "*tracking*" nas versões de software dos roteadores, fundamental para o projeto ser mais robusto e completo.

Os experimentos não foram conclusivos para avaliar qual o melhor protocolo para redundância, onde cada um tem suas particularidades e depende do cenário e equipamentos utilizados.

O trabalho deixou de abordar um experimento mais completo, abordando um cenário com vários tipos de falhas, avaliando conexões fim a fim, servindo para novas pesquisas e trabalhos futuros.

4 Experimento Integrando HSRP com Protocolo BGP

Neste capítulo será apresentado um experimento com a integração do protocolo de redundância avaliado no capitulo anterior o HSRP, com o protocolo de roteamento BGP. O objetivo é conseguir além de um sistema de redundância simulando falhas no sistema, o balanceamento de carga de tráfego de dados na estrutura disponível montada.

Com base num cenário real de uma empresa corporativa, será descrita na sessão 4.1 a topologia de rede de acesso atual da empresa, seguido na próxima sessão de um cenário pretendido para a execução do projeto e experimento.

Os experimentos serão executados com equipamentos reais em bancada. Suas aplicações e configurações serão apresentadas na sessão 4.3, e a execução do projeto com suas respectivas avaliações dos resultados obtidos na sessão 4.4.

4.1 Cenário Atual

O cenário atual é um modelo de uma rede *stub*, com apenas um CPE na rede local, e que possui um *link wan* ponto-a-ponto com velocidade de 10Mbps, através da tecnologia de acesso ATM fornecido por uma operadora para acessar uma rede remota externa e com acesso a internet, como mostra a figura 13.

O cenário não possui nenhum mecanismo de redundância. Todo o sistema da empresa depende da disponibilidade deste *link*. Ela se utiliza do *link* para acesso a banco de dados remoto e acesso a internet. Quedas no enlace são constantes e a tecnologia do *link wan* (ATM) é considerada uma rede legada. O roteamento é implementado de forma estática, e caso uma nova rede seja implementada, será necessário intervenção técnica manual.

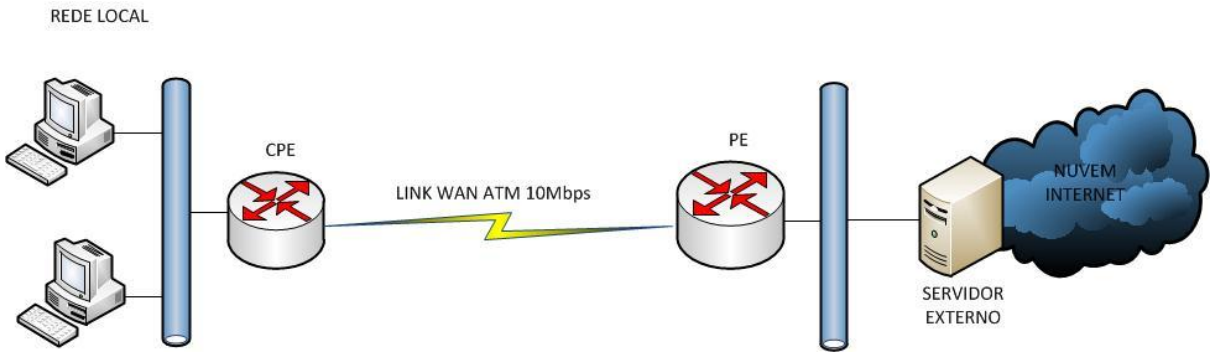

Figura 13: Topologia Atual da Empresa

4.2 Cenário Pretendido

A proposta do projeto foi implementar mais um circuito *wan* ponto-a-ponto, com dois CPE's na estrutura local da empresa, montando uma topologia de rede de múltiplos acessos e conectados a um PE conforme ilustrado na figura 14. Este cenário permite a utilização dos protocolos de redundância, neste caso foi escolhido explorar o protocolo HSRP conforme avaliações descritas no capítulo 3.

Além da utilização dos protocolos de duplicação de *gateway*, foi utilizado no sistema o protocolo de camada 3 integrando redundância e balanceamento entre os *links*, a fim de aproveitar toda a estrutura disponível. O BGP foi o protocolo de camada 3 explorado no projeto para configuração do balanceamento, utilizando-se de vários atributos para seleção de rotas (*AS path prepending*, *local-prefence*, *route-map*, rotas estáticas, etc).

O projete provê uma redundância de primeiro *hop* e do link de acesso *wan* de uma operadora com disponibilidade pós-queda dinamicamente, assegurando total disponibilidade de todas as aplicações da rede, porém com banda reduzida até a solução do problema. Podemos citar algumas falhas:

- Falha de link físico (operadora)
- Indisponibilidade do roteador (falta de energia)
- Queda de interface local(cabos de rede – switch)
- Perda de conectividade IP com roteador vizinho

O projeto não envolve disponibilidade de recursos de energia e redundância de rede local de switch e servidores.

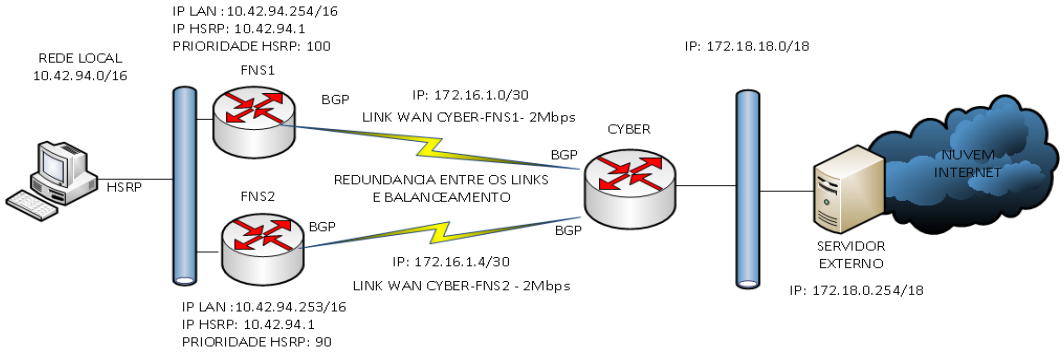

Figura 14. Topologia Proposta

As figuras a seguir ilustram de forma gráfica o balanceamento de tráfego de entrada e saída da rede que se espera com a proposta do projeto em três situações. Situação normal (figura 15), falha no enlace principal (figura 16) e falha no enlace redundante (Figura 17).

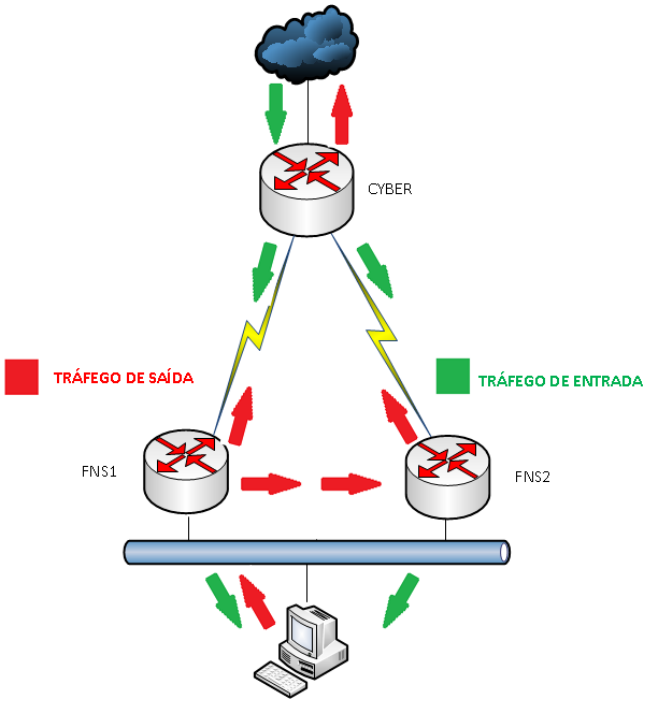

Figura 15: Situação Normal de tráfego

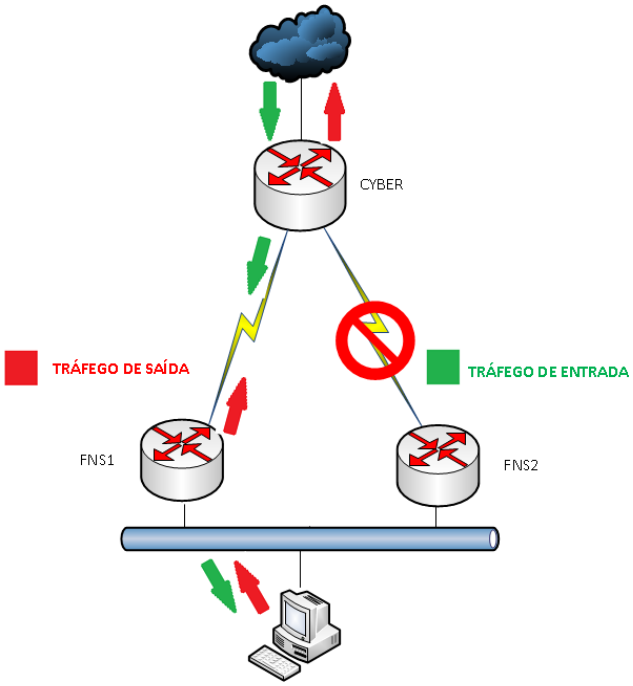

Figura 16: Situação de falha em FNS2

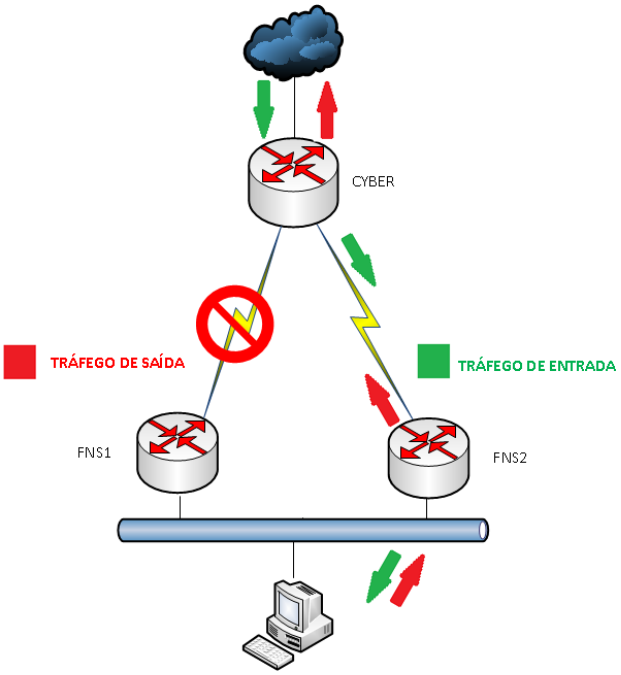

Figura 17: Situação de falha em FNS1

4.3 Descrição de configuração do experimento

O cenário foi montado em bancada conforme a topologia da figura 14 com 3 roteadores Cisco série 1700 denominados como: FNS1, FNS2 e CYBER, interligados por cabos serias V35 e enlace de modem, um switch layer 2 Encore e dois computadores. A velocidade dos enlaces é limitada em 2Mbps devido a utilização de interfaces seriais para a simulação, servindo apenas para o experimento, não sendo o ideal para o projeto.

No experimento foi aplicada uma rede *full-mesh,* onde todos os nós se comunicam entre si em forma de malha, com uma sessão BGP entre os roteadores vizinhos conforme a topologia. As sessões BGP foram configuradas com os ip's de *WAN* entre CYBER e FNS, e entre os dois roteadores na rede local com IP de LAN. As tabelas abaixo mostram as configurações do BGP de cada roteador com descrição dos comandos relevantes.

BGP – CYBER
router bgp 10 bgp log-neighbor-changes neighbor 172.16.1.2 remote-as 65000 – *Peer com FNS2* neighbor 172.16.1.6 remote-as 65000 – *Peer com FNS1* maximum-paths 2 maximum-paths ibgp 2 address-family ipv4 redistribute connected – *Comando para redistribuir suas redes conectadas* redistribute static – *Comando para redistribuir suas redes conectadas* neighbor 172.16.1.2 activate neighbor 172.16.1.6 activate maximum-paths 2 – *Comando para habilitar balanceamento* maximum-paths ibgp 2 - *Comando para habilitar balanceamento* default-information originate – *Comando para divulgar rota padrão (default).* no auto-summary no synchronization exit-address-family

Tabela 2: BGP CYBER

BGP – FNS1
router bgp 65000 bgp log-neighbor-changes neighbor 10.42.94.253 remote-as 65000 - *Peer com FNS2* neighbor 172.16.1.5 remote-as 10 – *Peer com CYBER* maximum-paths 2 maximum-paths ibgp 2 address-family ipv4 redistribute connected redistribute static neighbor 10.42.94.253 activate neighbor 172.16.1.5 activate maximum-paths 2 maximum-paths ibgp 2 no auto-summary no synchronization exit-address-family

Tabela 3: BGP FNS1

BGP – FNS2
router bgp 65000 bgp log-neighbor-changes neighbor 10.42.94.254 remote-as 65000 neighbor 172.16.1.1 remote-as 10 maximum-paths 2 address-family ipv4 redistribute connected redistribute static neighbor 10.42.94.254 activate neighbor 172.16.1.1 activate neighbor 172.16.1.1 route-map CORP in – *Mapeamento de rota de entrada* maximum-paths 2 default-information originate no auto-summary no synchronization network 10.42.94.0 mask 255.255.0.0 exit-address-family

Tabela 4: BGP FNS2

Com apenas um *gateway* ativo na rede, todo o tráfego de saída da rede local FNS será direcionado para o roteador FNS1, definido no HSRP como acesso principal da rede. O objetivo inicial é balancear o tráfego de saída entre os *links* com o CYBER e FNS2. Na configuração padrão do BGP, não foi possível o balanceamento, pois a tabela de roteamento montada em FNS1 sempre dá preferência para saída com o CYBER devido seu caminho ser mais curto. Um exemplo é para chegar no *gateway* do CYBER 172.18.0.254, o tráfego passaria por apenas um salto FNS1-CYBER, pelo link redundante passaria por dois saltos FNS1-FNS2-CYBER. A solução inicial foi à aplicação do atributo "*As-path Prepend*" na saída da sessão BGP entre os roteadores CYBER-FNS1, um atributo do BGP apresentado no trabalho, que pode ser aplicado para simular o incremento de saltos para o prefixos

recebidos, fazendo com que todas as rotas de FNS1 recebidas pelo CYBER tenham marcadas dois caminhos, “enganando” a tabela de roteamento, e igualando o custo das rotas.

Mesmo com esta solução aplicada, o roteador FNS1 não se comportou conforme o esperado balanceando tráfego entre seus vizinhos, dando preferência sempre para um caminho (FNS1-CYBER). Verificado que não foi possível o balanceamento desta forma devido a versão de software e hardware do roteador FNS1 não possuir a função de balanceamento por pacote nas suas interfaces (*ip load-sharing per-packet*).

Outra solução foi distribuir o tráfego no roteador FNS1 por destinos, determinando preferências por determinados prefixos de rede. No caso foi dividido o tráfego da rede em dois caminhos, um *link* para o tráfego de internet (prefixo rota padrão 0.0.0.0/0) e outro link para tráfegos da rede corporativa da empresa (prefixo 172.18.0.0/18). Nesta solução foi aplicada na entrada da tabela de roteamento do roteador FNS2, o atributo “*local-preference*” também relatado no trabalho, que tem a função de dar um valor de preferência maior que o padrão, no caso padrão é 100 e foi aplicado 200 apenas no prefixo 172.18.0.0/18 recebido pelo CYBER, distribuindo este prefixo também para o roteador FNS1. Para isso, teve que se aplicar um mapeamento de rota (*route-map*) na sessão BGP de entrada do FNS2 com CYBER, chamada de CORP. Por sua vez o roteador FNS2 divulga sua tabela de roteamento para o FNS1 com os parâmetros aplicados. Então, o roteador FNS1 recebe dois caminhos para a rede corporativa, um pelo CYBER com preferência padrão (100), e outra pelo FNS2 com preferência 200. Segue na figura 18 abaixo, a configuração de preferência no roteador FNS2 e na figura 19 a tabela de roteamento através do comando *show ip bgp*:

```
router bgp 65000
 bgp log-neighbor-changes
 neighbor 10.42.94.254 remote-as 65000
 neighbor 172.16.1.1 remote-as 10
 maximum-paths 2
 !
 address-family ipv4
 redistribute connected
 redistribute static
 neighbor 10.42.94.254 activate
 neighbor 172.16.1.1 activate
 neighbor 172.16.1.1 route-map CORP in
 maximum-paths 2
 default-information originate
 no auto-summary
 no synchronization
 network 10.42.94.0 mask 255.255.0.0
 exit-address-family
!
ip classless
no ip http server
!
!
!
ip prefix-list CORP seq 5 permit 172.18.0.0/18
!
route-map CORP permit 10
 match ip address prefix-list CORP
 set local-preference 200
!
route-map CORP permit 20
!
```

Figura 18: Seta Preferência 200 FNS2

```
FNS1#sh ip bgp 172.18.0.0
BGP routing table entry for 172.18.0.0/18, version 6
Paths: (2 available, best #2, table Default-IP-Routing-Table)
Multipath: eBGP iBGP
  Advertised to non peer-group peers:
  172.16.1.5
  10
    172.16.1.5 from 172.16.1.5 (172.18.18.98)
      Origin incomplete, metric 0, localpref 100, valid, external
  10
    172.16.1.1 from 10.42.94.253 (172.16.1.2)
      Origin incomplete, metric 0, localpref 200, valid, internal, best
```

Figura 19: Tabela de Roteamento

Percebe-se no retorno do comando dois caminhos para o prefixo 172.18.0.0/18. Um pelo link com CYBER (172.16.1.5) e outro pelo roteador FNS2 (10.42.94.253). Nota-se a preferência maior para o caminho FNS2 ("*localpref 200*"), e o roteador elegendo como melhor caminho ("*best*").

A solução para a rota padrão sair pelo *link* principal não foi preciso nenhum parâmetro adicional nas configurações do BGP, o roteador recebe duas rotas default também, porém a rota pelo link principal é definida pelo caminho mais curto. Segue tabela na figura 20.

```
FNS1#sh ip bgp 0.0.0.0
BGP routing table entry for 0.0.0.0/0, version 9
Paths: (2 available, best #1, table Default-IP-Routing-Table)
Multipath: eBGP iBGP
  Advertised to non peer-group peers:
  10.42.94.253
  10
    172.16.1.5 from 172.16.1.5 (172.18.18.98)
      Origin incomplete, metric 0, localpref 100, valid, external, best
  10
    172.16.1.1 from 10.42.94.253 (172.16.1.2)
      Origin incomplete, metric 0, localpref 100, valid, internal
```

Figura 20: Rota Padrão

Do lado do roteador CYBER, o roteador recebe através da sessão BGP o prefixo da rede local de FNS 10.42.94.0/16 pelos dois roteadores FNS1 e FNS2, balanceando o tráfego de saída para o sentido FNS. Diferente do roteador FNS1, foi possível o balanceamento através dos parâmetros aplicados no BGP *maximum-paths* 2 e *maximum-paths ibgp* 2 que habilita o roteador a ter um numero máximo de 2 caminhos para um mesmo prefixo, e o *ip load-sharing per-packet* nas interfaces *wan* habilitando roteamento entre as interfaces por pacotes. O retorno do comando "show ip route" para o prefixo 10.42.94.0/16 mostra os dois caminhos para se chegar ao destino, ilustrado na figura 21. No roteador CYBER, também foi aplicada uma rota padrão estática para o servidor 172.18.0.254, sendo distribuída através da

sessão BGP. Foi necessário o parâmetro *default-information originate* no BGP do CYBER para divulgação da rota padrão conforme tabela 2.

```
CYBER#sh ip route
Codes: C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route

Gateway of last resort is 172.18.0.254 to network 0.0.0.0

B        10.42.0.0 [20/0] via 172.16.1.6, 00:30:40
                  [20/0] via 172.16.1.2, 00:14:49
S*   0.0.0.0/0 [1/0] via 172.18.0.254
```

Figura 21: Show ip route CYBER

O HSRP foi configurado conforme a configuração mostrada na tabela 1. Porem foi adicionado um "*track*" na interface do roteador FNS1(*standby 1 track 1 decrement 30*). Este comando serve para o roteador usar como referencia para identificação da falha no seu acesso *wan*. Usualmente se utiliza o status das interface (*up/down*) nas aplicações do HSRP, porém nem sempre a falha de um acesso está ligada a problemas físicos. Para um sistema mais completo e robusto, foi utilizado a tabela de roteamento como parâmetro para queda do status do HSRP. Quando o *link* falhar, a sessão BGP também cai e consequentemente os prefixos de rede recebidos por aquela sessão BGP também cai. Quando o roteador identifica a queda do prefixo monitorado pela *track* 1 (*track 1 ip route 172.16.1.5 255.255.255.255 reachability*), o roteador decrementa sua prioridade em 30, no caso 100 – 30 = 70, abaixo do valor fixado em FNS2, valor 90. Nesta situação o protocolo HSRP envia mensagens "*hello*" com a nova prioridade do roteador FNS1 em *multicast ethernet*. O roteador FNS2 identifica a mensagem, muda seu status de *stand-by* para *active*, fazendo com que o roteador FNS2 assuma como gateway ativo da rede.

4.4 Execução e avaliação dos resultados

Os testes e simulações de tráfego foram feitos através de aplicativos geradores de tráfego UDP/TCP JPERF, *downloads* da internet e transferência de arquivos FTP de um servidor local e externo. Percebeu-se que foi possível utilizar os dois links funcionando conforme o esperado, a saída FNS1-CYBER com tráfego para internet, e FNS-CYBER para rede corporativa. O retorno do tráfego se apresentou balanceado entre os links devido ao balanceamento por pacote. Este tipo de balanceamento pode ser considerado um problema em aplicações não orientado a conexão (UDP), como voip e multimídia, principalmente se os links forem assimétricos. Porém, não foram testados no experimento este tipo de aplicação.

A solução em conjunto com o protocolo de duplicidade de gateway HSRP, com protocolo de roteamento BGP definida no experimento para o balanceamento de tráfego entre os *links* de acesso *wan*, permite também a solução de redundância na falha de um dos *links*. É esperado que o tráfego de internet migre para o roteador FNS2 quando o acesso FNS1-CYBER falhar, e FNS2 assumindo todo o tráfego da rede, e vice e versa. Os testes serão analisados simulando a queda dois links alternadamente. Os dados do experimento serão coletados com aplicativos de gerenciamento de tráfego das interfaces dos roteadores através de gráficos e logs dos roteadores.

Para testar a redundância dos links, foi gerado com o aplicativo JPERF um tráfego limitado em 1Mbps para uma melhor ilustração da rede FNS para um ip da Internet (8.8.8.8) e para um host da rede corporativa (172.18.18.98) ilustrado na figura 22.

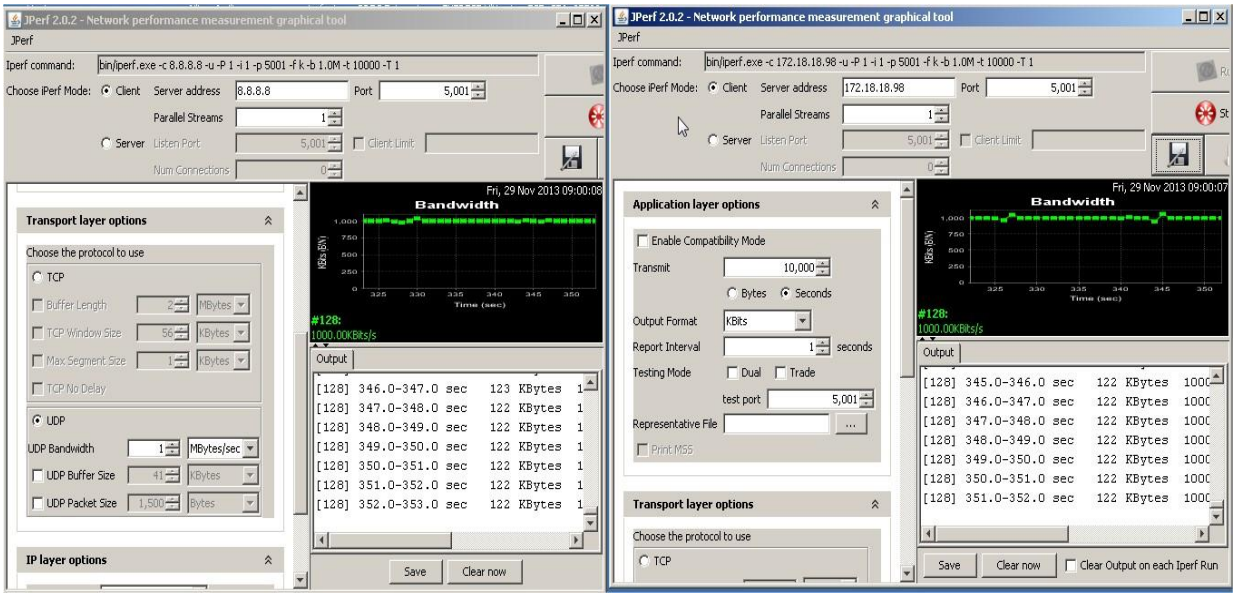

Figura 22: Jperf

Os roteadores assumiram o tráfego conforme o esperado: tráfego de internet entre o *link* FNS1-CYBER e corporativo entre FNS2-CYBER. Com testes de *traceroute* padrão do Windows a partir do host de FNS, é possível perceber os saltos para os determinados destinos conforme figuras 23 e 24.

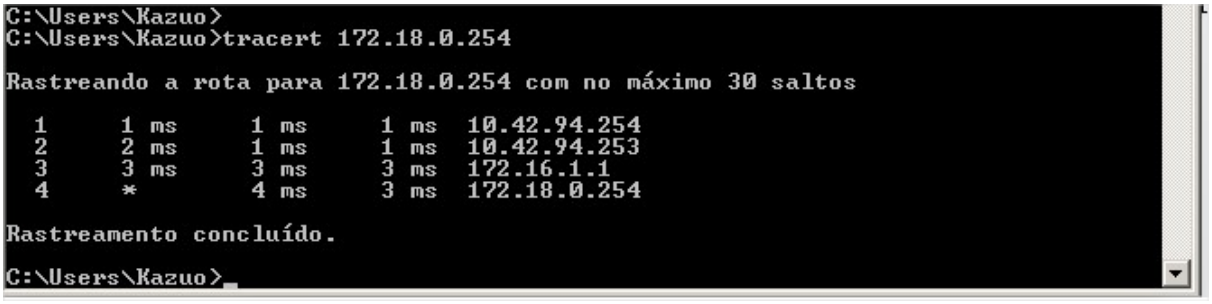

Figura 23: Tracert rede corporativa

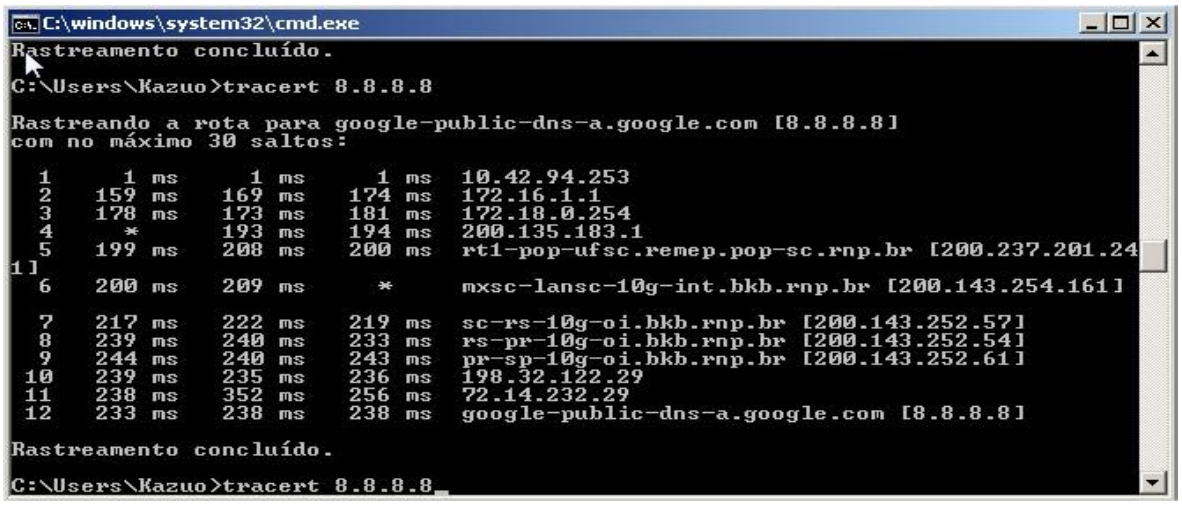

Figura 24: Tracert Internet

Percebe-se que na figura 24 do *traceroute* para a rede corporativa, o pacote a partir do host foi enviado para o seu gateway ativo FNS1 10.42.94.254, enviado para o próximo salto para o roteador FNS2 10.42.94.253, passando pelo link WAN até chegar no seu destino.

Para ilustrar a solução de balanceamento, foi utilizado um aplicativo de monitoramento de rede PRTG, a fim de gerar gráficos das interfaces dos roteadores. Abaixo o gráfico 1 ilustra através do tráfego gerados pelo JPERF a migração do tráfego de dados entre os acessos com as simulações de queda nos circuitos.

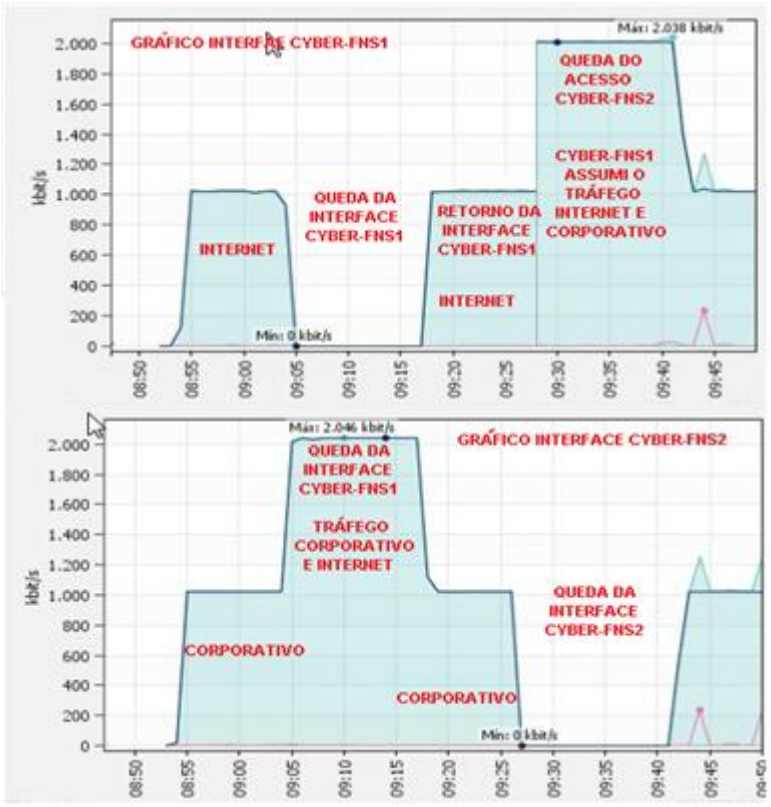

Gráfico 1: Balanceamento de Carga

Percebe-se no gráfico o comportamento esperado pelas simulações. Os tráfegos gerados pelo aplicativo JPERF de um 1Mbps para cada acesso, as 09:05 se simula a queda do acesso CYBER-FNS1, fazendo com que CYBER-FNS2 assuma todo o tráfego chegando em 2Mbps. Ao retornar o acesso em 09:17, o tráfego volta ao seu estado original. Consequentemente as 09:27, se simula a queda da interface de FNS-CYBER, fazendo que todo tráfego da rede corporativa migre para o link CYBER-FNS1. Ao normalizar os acessos, o sistema e a distribuição do tráfego volta ao seu estado original.

As figuras a seguir, mostram de forma ilustrativa os sentidos dos tráfegos conseguidos com o experimento, separando o tráfego de internet e corporativo em três situações de funcionamento. A figura 25 mostra o comportamento do tráfego com o sistema sem falhas, figura 26 com a queda do link CYBER-FNS1, e a figura 27 com o *link* CYBER-FNS2 indisponível.

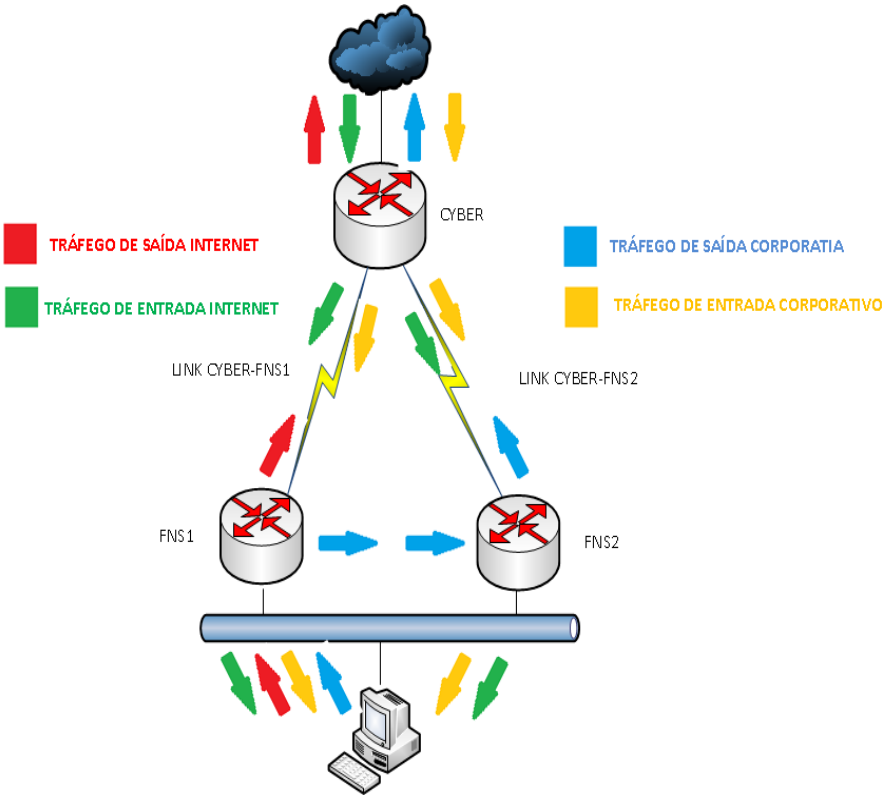

Figura 25: Roteamento sem Falhas

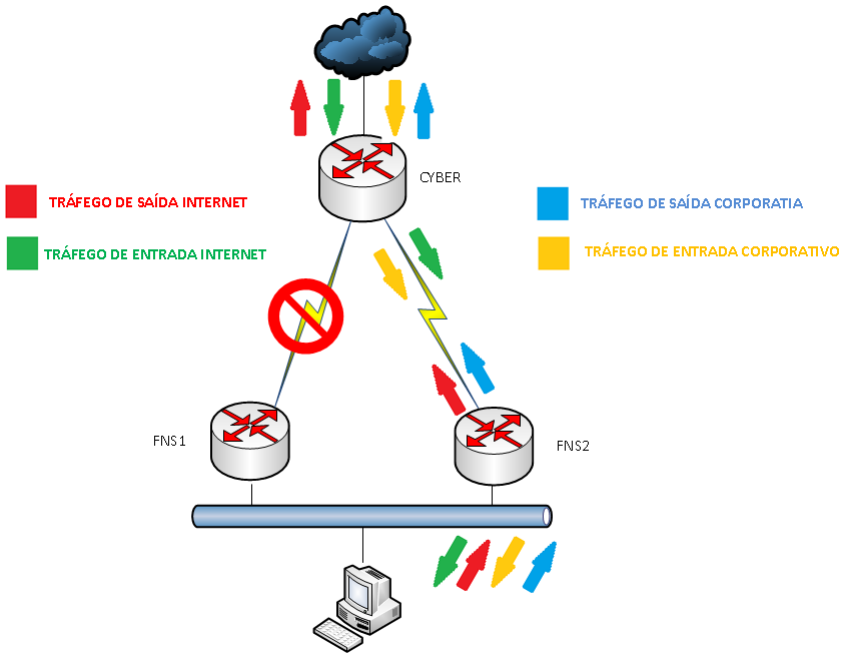

Figura 26: Roteamento CYBER-FNS1 em falha

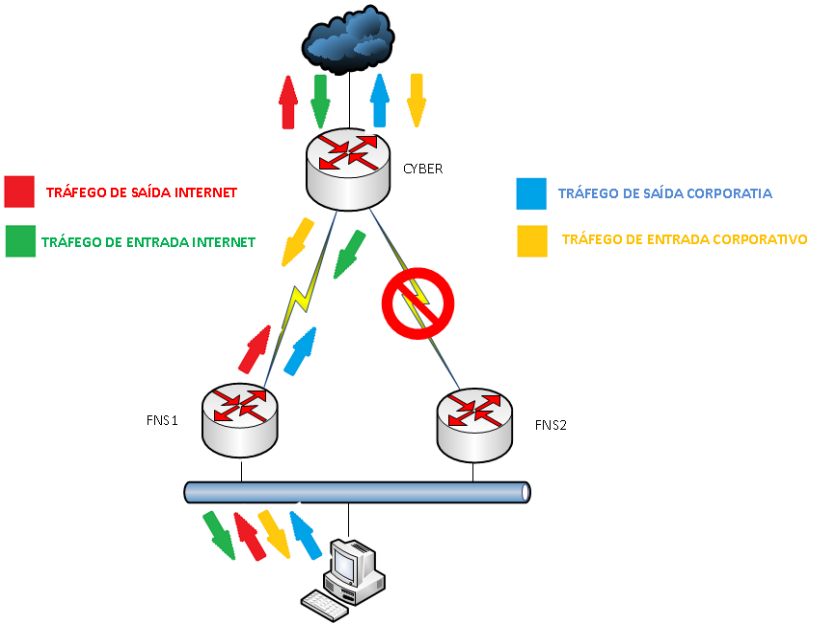

Figura 27: Roteamento CYBER-FNS2 em falha

Abaixo a figura 28 ilustra um exemplo dos logs do roteador FNS1 quando simulado a queda de sua interface. Verifica-se que as 11:06:33 cai a interface wan Serial 0/0, consequentemente o protocolo BGP, o HSRP percebe a queda de sua rota, decrementa o sua prioridade a passa seu status para “Speak”. As 11:13 a interface volta a subir e o HSRP volta ao seu estado de ativo.

```
*Mar  7 11:06:33.953: %LINK-3-UPDOWN: Interface Serial0/0, changed state to down
*Mar  7 11:06:33.969: %BGP-5-ADJCHANGE: neighbor 172.16.1.5 Down Interface flap
*Mar  7 11:06:34.953: %LINEPROTO-5-UPDOWN: Line protocol on Interface Serial0/0, changed state to down
*Mar  7 11:06:34.961: %HSRP-6-STATECHANGE: FastEthernet0/0 Grp 1 state Active -> Speak
*Mar  7 11:13:09.205: %LINK-3-UPDOWN: Interface Serial0/0, changed state to up
*Mar  7 11:13:12.217: %LINEPROTO-5-UPDOWN: Line protocol on Interface Serial0/0, changed state to up
*Mar  7 11:13:15.797: %BGP-5-ADJCHANGE: neighbor 172.16.1.5 Up
*Mar  7 11:13:17.049: %HSRP-6-STATECHANGE: FastEthernet0/0 Grp 1 state Standby -> Active
```

Figura 28: Log FNS1

Nos testes de traceroute padrão do Windows, segue a variação dos caminhos percorridos no exemplo da falha simulada no link CYBER-FNS1. Inicialmente o traceroute percorre seu caminho original, FNS1(10.42.94.254), saindo pela acesso CYBER-FNS1(172.16.1.5), e saindo pela internet através do CYBER (172.18.0.254), figura 29 . Após a simulação de queda, o caminho do tráfego muda para o seu novo gateway padrão FNS2 (10.42.94.253) ilustrado na figura 30.

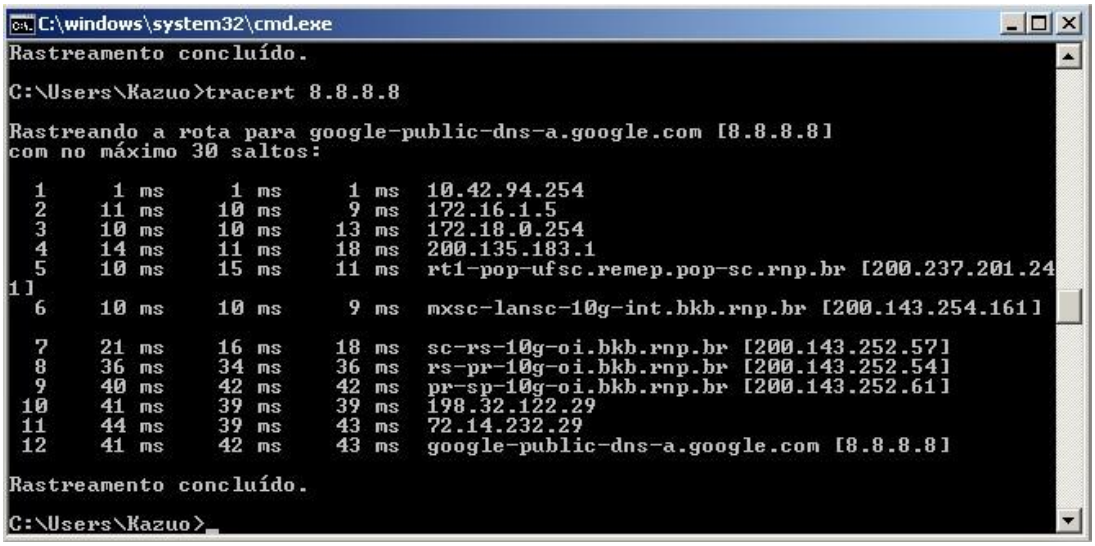

Figura 29: Traceroute Internet normal

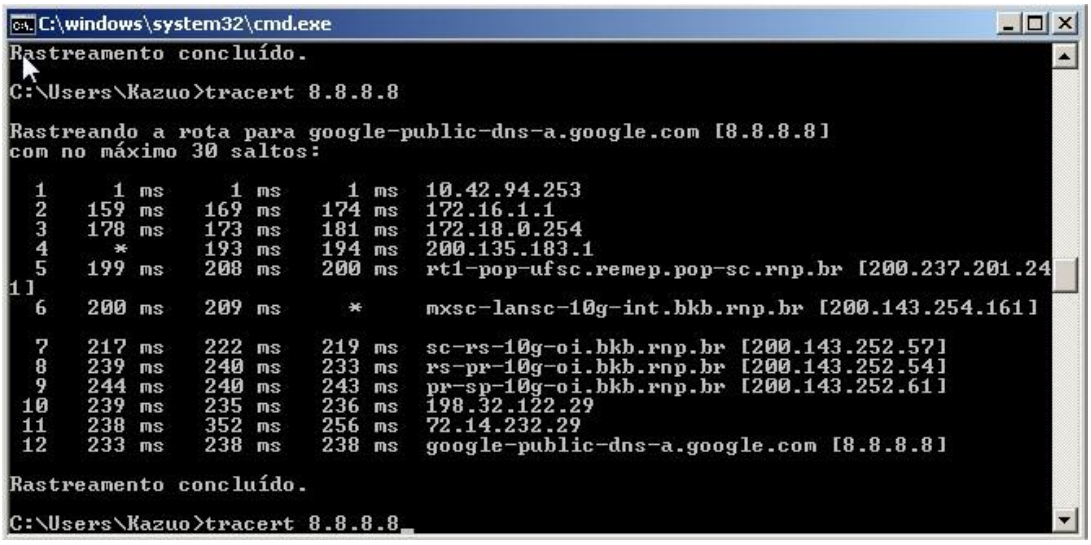

Figura 30: Traceroute Internet com redundância

4.5 Conclusão

Nos experimentos e testes simulados conforme descrito neste capitulo, por problemas de software ou hardware dos roteadores não foi conseguido o objetivo inicial de balancear o tráfego da rede entre os dois links de acesso disponíveis na topologia, de forma balanceada e simétrica. Com a aplicação de alguns atributos no protocolo BGP, a solução utilizada foi dividir o tráfego da rede nos dois acessos utilizando o destino do tráfego como parâmetro.

Foi divido o tráfego de internet para o link CYBER-FNS1 e o tráfego corporativo entre CYBER-FNS2, conseguindo então utilizar os dois acessos simultaneamente. A solução de redundância funcionou conforme o esperado, e além de alterar o gateway da rede para link do

roteador sem falha, os dois tráfegos são migrados para o acesso sem falha, provendo um sistema de redundância mais completo, aproveitado de todo o sistema de acesso e provendo maior banda para cada tráfego.

5 Conclusão

Este trabalho de conclusão de curso teve como objetivo mostrar, projetar, experimentar e avaliar o uso combinado de protocolos de redundância de primeiro hop com protocolos de roteamento, a fim de conseguir um sistema de rede de acesso com redundância e balanceamento de carga, utilizando-se de um cenário real de uma empresa corporativa.

Na primeira etapa do trabalho, foi feito pesquisas com o objetivo de aprofundar o tema: *Redundância e Balanceamento em Rede Corporativa* e ter embasamento teórico para realizar análises e definir a proposta do trabalho.

Os estudos mostraram que no cenário em que foi abordado o projeto, além da redundância que foi implementada, pode-se também aplicar balanceamento de carga no sistema se utilizado em conjunto com protocolos de roteamento em camada 3. Para isso, foi feito experimentos com equipamentos reais em bancada para avaliar o desempenho dos protocolos de redundância. Após análise, um novo experimento foi feito simulando um cenário real de uma empresa, integrando o protocolo de redundância HRSP com protocolo de roteamento BGP, a fim de obter balanceamento de carga.

Com os testes e simulações realizadas, a proposta inicial de integrar os dois protocolos para prover redundância e balanceamento de carga foi alcançada. Porém, o objetivo de balancear o tráfego de forma simétrica não foi almejado por limitações dos equipamentos. Com o experimento foi conseguido um sistema de redundância e distribuição de tráfego entre os acessos, dividindo a rede em prefixos utilizando de todo o recurso disponível, provendo mais banda disponível e sem um recurso ocioso desperdiçado.

O experimento final foi limitado utilizando apenas o protocolo de redundância HSRP e de roteamento BGP. Para trabalhos futuros, este trabalho pode servir como base para integração com outros protocolos de roteamento, como exemplo: OSPF, RIP dentre outros. Pode-se também avaliar com os demais protocolos de duplicação de gateway VRRP e GLBP. O projeto também não explorou outros recursos de redundância como: fontes de energia, redundância em *layer* 2 (switches redundantes), duplicidade de servidores com VRRP e demais.

Provou-se que é possível ter um sistema redundante com um melhor aproveitamento e que este recurso pode ser explorado por operadoras em clientes corporativos. O trabalho também teve o intuito de mostrar recursos e protocolos que obtive conhecimento no meu âmbito profissional, e que não são lecionados no currículo deste curso.

Referências Bibliográficas

AVIZIENIS, A. et al, "Basic Concepts and Taxonomy of Dependable and Secure Computing," IEEE Transactions on Dependable and Secure Computing, vol. 1, no. 1, pp. 11-33, January-March, 2004.

HINDEN. “RFC 3768: Virtual Router Redundancy Protocol (VRRP)", Abril 2004. Disponível em: <http://www.ietf.org/rfc/rfc3768.txt>. Acesso em: 29 jan. 2013.

ISEL. Protocolo BGP. . Transparências ISEL http://www.deetc.isel.ipl.pt/redesdecomunic/disciplinas/RI/acetatos/ProtocoloBGP.pdf. Acessado em 7 de março de 2013.

KRAEMER, VILAR, GOLDMAN. Tolerância a Falhas utilizando protocolos de *Gateway* Redundantes. Disponível em: http://www.ime.usp.br/~gold/publications/pdf/erad2010.pdf. Acesso em: 29 jan. 2013.

MEDHI, DEEPANKAR; RAMASAMY, KARTHIKEYAN. Network Routing Algorithms, Protocols, and Architectures. Morgan Kaufmann Publishers . Elsevier, 2007.

PINHEIRO, JOSÉ MAURÍCIO DOS SANTOS. Conceitos de Redudância e Contingência.http://www.projetoderedes.com.br/artigos/artigo_conceitos_de_redundancia.php#.UTjFnxJYsuJ. Acessado em 7 de março de 2013.

RISSO, FULVIO. Redundancy and load balancing at L3 in Local Area Networks Politecnico di Torino .http://netgroup.polito.it/teaching/prlc/LAN%20-%20L3%20redundancy.pdf Acessado em 7 de março de 2013.

Anexos

Anexo 1:

HSRP	hr da queda	hr ativo	tempo	pacotes perdidos
teste 1	08:01:57.343	08:01:59.154	2	2
teste 2	08:03:56.234	08:03:57.977	1	2
teste 3	08:06:16.787	08:06:17.409	1	2
teste 4	08:08:47.123	08:08:53.131	6	3
teste 5	08:13:17.541	08:13:22.231	5	2
teste 6	08:15:37.876	08:15:41.786	4	2
teste 7	08:18:07.275	08:18:09.423	2	2
teste 8	08:18:57.980	08:19:04.987	7	2
teste 9	08:20:27.123	08:20:29.276	2	2
teste 10	08:21:37.765	08:21:33.137	6	2
Média			3,6	2,1
VRRP	hr da queda	hr ativo	tempo	pacotes perdidos
teste 1	08:30:27.691	08:30:28.271	1	1
teste 2	08:35:02.159	08:35:03.159	1	1
teste 3	08:38:38.251	08:38:37.387	1	1
teste 4	08:41:40.967	08:41:41.927	1	1
teste 5	08:43:34.531	08:43:35.451	1	1
teste 6	08:44:45.683	08:44:46.831	1	1
teste 7	08:45:59.799	08:46:01.411	2	1
teste 8	08:47:24.843	08:47:26.247	2	1
teste 9	08:48:23.855	08:48:25.223	2	1
teste 10	08:49:38.855	08:49:40.483	2	1
Média			1,4	1
GLBP	hr da queda	hr ativo	tempo	pacotes perdidos
teste 1	08:56:47.619	08:56:54.619	7	0
teste 2	09:00:59.527	09:01:04.631	5	2
teste 3	09:02:55.595	09:03:02.647	7	1
teste 4	09:04:16.175	09:04:21.647	5	0
teste 5	09:05:49.711	09:05:55.647	6	1
teste 6	09:08:08.563	09:08:14.643	6	2
teste 7	09:10:40.215	09:10:45.643	5	2
teste 8	09:13:46.167	09:13:52.643	6	1
teste 9	09:15:18.583	09:15:23.643	5	0
teste 10	09:18:17.303	09:18:21.639	4	0
Média			5,6	0,9