Abordagem para Desviar Pacotes na Presença de Falha em Backbones IP com OSPF

Fernando Barreto[1], Emilio C. G. Wille[1], Luiz Nacamura Junior[1]

[1]Pós-Graduação em Engenharia Elétrica e Informática Industrial – Universidade Tecnológica Federal do Paraná (UTFPR)
CEP – 80230-901 – Curitiba – PR – Brazil

{fbarreto,ewille,nacamura}@utfpr.edu.br

Abstract. *IP network backbones use link state routing protocols to find correct routes. In face of a topology change, e.g. a failure, these protocols need some time to react to it in order to find new routes. During this time, the routes become unstable, causing high packet loss rate and depreciation of backbone reliability. This work presents a proactive approach to help the IP routing protocols during the convergence period in order to reduce packet loss rate. The approach is evaluated using various artificial and real topologies, and a simulation is also implemented in order to evaluate the packet loss rate reduction, which yielded satisfactory results.*

Resumo. *Os backbones IP utilizam protocolos de roteamento do tipo estado do enlace para definir as rotas corretamente. Em situações de mudança na topologia, como uma falha, esses protocolos necessitam de um tempo para reagir e encontrar novas rotas. Durante esse tempo, as rotas ficam instáveis com alta taxa de pacotes perdidos e queda na confiabilidade do backbone. Esse trabalho propõe uma abordagem pró-ativa para auxiliar os protocolos de roteamento a reduzir a taxa de pacotes perdidos durante esse período. Essa abordagem é avaliada em representações de topologias artificiais e reais, e também em simulação para analisar a redução de pacotes perdidos, demonstrando resultados bastante satisfatórios.*

1. Introdução

As infra-estruturas de redes IP atuais são utilizadas como ambiente de transporte para tráfego de informações em geral. Devido ao seu uso genérico e crescente, tais infra-estruturas se tornam críticas e necessitam de alta confiabilidade, pois a interrupção de serviço causada por uma falha pode comprometer o desempenho dos tráfegos em geral. Dentre os cenários de falha existentes, as falhas transitórias e de um único componente de rede são as de maior ocorrência em *backbones* IP [Iannaccone et al. 2004] [Nelakuditi et al. 2007]. Entende-se por falha de um único componente de rede como sendo a falha de um enlace, roteador ou *Shared Risk Link Group* (SRLG). Um SRLG pode ser considerado um único componente, pois é constituído de um conjunto de enlaces dependentes que pertencem a um mesmo meio comum, e.g. duto com várias fibras, onde a falha de um enlace deve implicar na falha dos demais enlaces dependentes.

O processo de recuperação de uma falha deve ser o mais rápido possível para reduzir significativamente o comprometimento dos fluxos de tráfego passantes. O objetivo é manter

na medida do possível o nível de confiabilidade da infra-estrutura de rede. Dentre as abordagens utilizadas para esse processo de recuperação tem-se a recuperação em nível de rede e em nível de *hardware*. Em nível de rede tem-se a recuperação por recálculo de rotas [Iannaccone et al. 2004] e a proteção por rotas alternativas pré-calculadas [Shand and Bryant 2007]. Já em nível de *hardware* tem-se a instalação e distribuição de caminhos físicos de proteção (mais eficiente, porém muito mais cara).

O processo de recuperação por recálculo de rotas é um processo reativo executado pelos protocolos de roteamento atuais do tipo estado do enlace, como o OSPF [Moy 1998]. Para que o OSPF realize esse processo, ele deve reconhecer a falha, notificar a ocorrência aos roteadores vizinhos (que por sua vez divulgam para os seus vizinhos até cobrir toda a topologia), esperar um tempo necessário para receber todas as notificações dos outros roteadores, calcular os menores caminhos (SPF – algoritmo de Dijkstra) com as novas informações do estado do enlace, para então atualizar a tabela de encaminhamento (*Forwarding Information Base* – FIB). Todo esse processo, nomeado período de convergência, atinge em torno de dezenas de segundos [Iannaccone et al. 2004], o que compromete seriamente a confiabilidade da infra-estrutura de rede. Para reduzir esse período, [Francois et al. 2005] modificou o tempo de cada item integrante desse processo tornando possível reduzir esse período para a casa de décimos de segundo. Porém, essa redução pode causar mais instabilidades [Basu and Riecke 2001] sendo dependente da topologia, da velocidade de processamento dos roteadores e dos enlaces.

As rotas alternativas pré-calculadas são abordagens que seguem o *framework* de *IP Fast ReRouting* (IPFRR) [Shand and Bryant 2007]. Esse *framework* define um esquema teórico para realizar um processo de recuperação local com caminhos previamente calculados para contornar uma falha. Dentre as principais abordagens IPFRR tem-se: FIR [Nelakuditi et al. 2007], NotVia [Bryant et al. 2007], LFA [Atlas and Zinin 2007], Tunnels [Bryan et al. 2005] e MRC [Kvalbein et al. 2006]. Entretanto, apenas o NotVia consegue atingir 100% de cobertura para falha de um enlace, roteador ou SRLG. O NotVia distribui endereços *not-via* na topologia para representar cada componente de rede a ser contornado. Uma rota para um endereço *not-via* em todos os roteadores consegue definir um caminho que contorna o componente que ele representa. Entretanto, o NotVia pode gerar caminhos mais extensos que o necessário para desviar os pacotes devido à localização dos endereços *not-via* (*next-next-hop*), além de ocupar recursos excessivos na FIB para representá-los e depender da técnica de encapsulamento (como conseqüência pode fragmentar ou até descartar os pacotes). Maiores detalhes sobre o NotVia e as demais abordagens não são descritos por restrição de espaço.

Este trabalho propõe uma nova abordagem formalizada e completa da proposta inicial do *Esquema de Caminhos Emergenciais Rápidos* (E-CER) [Barreto et al. 2007], incluindo a extensão na FIB. Essa abordagem adapta a idéia do *framework* IPFRR para fornecer caminhos mais curtos que as abordagens IPFRR e, dependendo da falha, sempre iguais aos obtidos pelo processo OSPF após sua convergência (permite uma adaptação gradual às novas rotas OSPF). Além disso, o E-CER necessita de menos recursos para representar esses caminhos na FIB e independe da técnica de encapsulamento.

2. E-CER

E-CER é uma abordagem distribuída baseada no *framework* IPFRR para calcular os

caminhos de recuperação, cada um nomeado *Caminho Emergencial Rápido* (CER). Esses caminhos são disponibilizados em forma de marcas (*CER_Marca/IR*) na FIB, que são utilizadas por um processo auxiliar de encaminhamento nomeado *CER_EncDif*.

Para uma abordagem de recuperação de falha cobrir 100% da topologia, esta deve possuir uma infra-estrutura física capaz de manter a conectividade entre todos os roteadores na presença de uma falha de enlace, roteador ou SRLG. Além disso, uma distribuição do tráfego de rede na topologia deve ser planejada para ocupar, no máximo, 50% da capacidade dos enlaces. Este planejamento já é realizado em várias topologias reais para permitir a acomodação de tráfego desviado na presença de falhas [Iannaccone et al. 2004]. Todas essas restrições podem ser adotadas por qualquer abordagem IPFRR, pois caso contrário, não será possível alcançar 100% de cobertura de falha.

2.1. Cálculo dos CERs

O E-CER considera a métrica OSPF existente (soma dos custos dos enlaces [Moy 1998]) e o número de roteadores na geração dos CERs. O uso da métrica do OSPF permite o reuso das rotas já existentes, o que reduz a complexidade do cálculo desses caminhos. Este trabalho adapta outras abordagens IPFRR para melhor utilizá-las, suporta SRLG, pesos de enlaces assimétricos, independe da técnica de encapsulamento e atinge 100% de cobertura de falha.

Cada roteador gera seus próprios CERs e este cálculo deve seguir a execução do processo OSPF, pois tem por objetivo reusar a base de dados do estado do enlace atualizada com a configuração da topologia.

Considere uma topologia de rede representada em um grafo G(*V*,*A*), onde *V* é o conjunto de roteadores e *A* o conjunto de enlaces que conecta os roteadores. Considere $(i, j) \in A$ a representação de um enlace conectando os roteadores *i* e *j*. Para cada (i, j), um custo $c_{i,j}$ é especificado durante a configuração do OSPF.

O cálculo SPF executado por um *Roteador Origem* (*RO*) basicamente minimiza a soma dos custos dos enlaces $c_{i,j}$ de *RO* até qualquer outro roteador na topologia. A primeira formulação matemática para cálculo dos CERs é uma adaptação da formulação de programação linear para o *Shortest Path Problem* [Ahuja et al. 1993]. O cálculo localiza a função custo $Z_{RO,J,RA}$ que minimiza a soma dos custos dos enlaces respeitando o número de unidades de fluxo que atravessam os enlaces para descobrir os caminhos de *RO* até um conjunto de roteadores destino *J*, onde $RO \notin J$, e impede o uso dos componentes adjacentes ao *RO*: *Roteador Adjacente* (*RA*) ou *Enlace Adjacente (RO,RA)*.

A restrição *(a)* seleciona os menores caminhos transmitindo *n-1* unidades de fluxo de *RO*, significando que a cada roteador alcançado no menor caminho 1 unidade de fluxo é consumida. As restrições *(b)* e *(c)* impedem os caminhos que utilizam *(RO,RA)*, se *J=RA*, ou *RA* (todos os enlaces adjacentes a *RA*), se *J≠RA*, pois define um custo de enlace infinito. Dessa forma, a formulação procura sempre encontrar caminhos que contornam *RA* quando possível, pois *RO*, na presença de falha, não consegue identificar imediatamente se uma falha é do *(RO,RA)* ou do *RA* e o contorno do *RA* automaticamente propicia o contorno do *(RO,RA)*. Se existem enlaces pertencentes a um SRLG que contém *(RO,RA), se J=RA*, ou *(RA,?)*, se *J≠RA*, a restrição *(d)* atribui pesos infinitos a esses enlaces para serem evitados na escolha do menor caminho.

$$Z_{RO,J,RA} = \min \sum_{i=1}^{n} \sum_{j=1}^{n} c_{i,j} x_{i,j}$$

Sujeito a:

$$\sum_{j=1}^{n} x_{i,j} - \sum_{k=1}^{n} x_{k,i} = \begin{cases} n\text{-}1, & if\ i = 1 \\ -1, & if\ i \neq 1 \end{cases} \quad (a)$$

$$c_{RO,j} = \infty \ \ se\ \ j = RA\ e\ se\ \ J = RA \quad (b)$$

$$c_{i,j} = \infty \ \ se\ j = RA\ e\ se\ \ J \neq RA \quad (c)$$

$$c_{i,j} = \infty \ \ se\ (i,j) \in \begin{cases} SRLG(RO,RA)\ ,\ se\ J = RA \\ SRLG(RA,?),\ se\ J \neq RA \end{cases} \quad (d)$$

$$i > 0\ ;\ j > 0\ ;\ i \neq j\ \ ;\ \ x_{i,j} \geq 0\ ;\ c_{i,j} \geq 0 \quad (e)$$

A solução para esse problema pode ser obtida com uma modificação simples no algoritmo de Dijkstra (SPF) para considerar as restrições *(b)*, *(c)* e *(d)*. Como apenas parte da árvore SPF é afetada pelo contorno do componente, o tempo de recálculo dos menores caminhos com *Incremental-SPF* [Narvaez 2000] é reduzido consideravelmente.

Os menores caminhos que obedecem a $Z_{RO,J,RA}$ são representados em um conjunto $\varphi_{RO,J,RA}$. Considere os menores caminhos para um *Roteador Destino* (*RD*) que pertencem ao $\varphi_{RO,J,RA}$, onde $RD \in J$. Se *RD* tem o seu menor caminho original OSPF (com todos os componentes na topologia) de *RO* até *RD* utilizando *(RO,RA)*, então os menores caminhos obtidos com $Z_{RO,J,RA}$ são *caminhos alternativos* para *RD* e são candidatos para contornar *(RO,RA)* ou *RA* dependendo se *RD=RA*. O conjunto de possíveis caminhos alternativos para *RD* que possui essa característica são representados como $\varphi_{RO,RD,RA}$. Cada caminho alternativo pertencente a $\varphi_{RO,RD,RA}$ é representado como $ALT_{RO,RD,RA}$.

A segunda formulação matemática abaixo ($Z'_{RO,RD,RA}$) foi criada para identificar, em cada $ALT_{RO,RD,RA}$, qual roteador é capaz de contornar uma falha de componente, i.e. *(RO,RA)* ou *RA*, quando, desse roteador, o caminho original OSPF pode seguramente encaminhar os pacotes pelo IP de destino até *RD*. O processo de identificação adapta a proposta do IPFRR [Shand and Bryant 2007] para criar uma classificação de níveis (ECMP, LFA e SIG) e facilitar a localização desse roteador, que recebe o nome de *Roteador CER* (*RC*). Cada roteador de $ALT_{RO,RD,RA}$ que é analisado na formulação para ser *RC* é nomeado *k*. O nível ECMP identifica o *RC* no primeiro roteador (*Roteador Vizinho – RV*) de $ALT_{RO,RD,RA}$ e o $ALT_{RO,RD,RA}$ deve ter o mesmo custo OSPF (soma dos custos dos enlaces, i.e. *DistOSPF*) que o caminho original OSPF de *RO* até *RD*. O nível LFA adapta a abordagem LFA original [Atlas and Zinin 2007] para identificar *RC* no *RV* e ainda considera a restrição do número de roteadores no caminho alternativo. O nível SIG é uma extensão da abordagem LFA para identificar o *RC* após o *RV*, dessa forma *k* segue através da seqüência de roteadores de $ALT_{RO,RD,RA}$ até encontrar um roteador capaz de ser o *RC*. Todos os *k* que não são capazes de ser *RC* são definidos como *roteadores_intermediários*. Portanto, independente do nível utilizado, um sub-caminho é gerado para cada $ALT_{RO,RD,RA}$ e recebe a identificação de $CER_{RO,RD,RA}$. Um $CER_{RO,RD,RA}$ é uma seqüência de roteadores até *RC*: se nível ECMP ou LFA então a seqüência é somente {*RC*} com *RC=RV*, já no nível SIG a seqüência é {*RV*, *roteadores_intermediários, RC*}. Todos os $CER_{RO,RD,RA}$ gerados para $\varphi_{RO,RD,RA}$ são candidatos ao CER selecionado para o *RD*, nomeado $S\_CER_{RO,RD,RA}$.

$$Z'_{RO,RD,RA} = \min(custo\_CER_{RO,RD,RA}\ num\_CER_{RO,RD,RA})$$

Sujeito a:

$$\varphi_{RO,RD,RA} \subset \varphi_{RO,J,RA} \quad \text{(a)}$$

$$ALT_{RO,RD,RA} \in \varphi_{RO,RD,RA} \quad \text{(b)}$$

$$CER_{RO,RD,RA} \subset ALT_{RO,RD,RA} \quad \text{(c)}$$

$$RC = k,\ se\ DistOSPF(RO,k)+DistOSPF(k,RD)=DistOSPF(RO,RA)+DistOSPF(RA,RD),\ e\ \ k \in Vizinhos\_RO \quad \text{(d)}$$

$$RC = k,\ se\ DistOSPF(k,RD)<DistOSPF(k,RO)+DistOSPF(RO,RD),\ e\ \ k \in Vizinhos\_RO \quad \text{(e)}$$

$$RC = k,\ se\ DistOSPF(k,RD)<DistOSPF(k,RO)+DistOSPF(RO,RD),\ e\ \ k \notin Vizinhos\_RO \quad \text{(f)}$$

$$RA \notin OSPF\_Roteadores(k,RD),\ se\ RA \neq RD \quad \text{(g)}$$

$$OSPF\_Enlaces(k,RD) \cap \begin{cases} SRLG(RO,RA)=\varnothing,\ se\ RA=RD \\ SRLG(RA,?)=\varnothing,\ se\ RA \neq RD \end{cases} \quad \text{(h)}$$

$$k \in CER_{RO,RD,RA},\ se\ nenhuma\ restrição\ (d),\ (e)\ ou\ (f)\ for\ possível \quad \text{(i)}$$

$$k \in CER_{RO,RD,RA}\ e\ k\ é\ o\ último\ roteador,\ se\ (RC=k)\ e\ (uma\ restrição\ (d),(e)\ ou\ (f)\ for\ possível) \quad \text{(j)}$$

$$num\_CER_{RO,RD,RA} = \begin{cases} 1\ unidade,\ se\ restrição\ (d)\ for\ possível \\ NumRoteadoresOSPF(k,RD),\ se\ restrição\ (e)\ for\ possível \\ NumRoteadoresCER_{RO,RD,RA},\ se\ restrição\ (f)\ for\ possível \end{cases}, \quad se\ (j)\ for\ possível \quad \text{(k)}$$

$$custo\_CER_{RO,RD,RA} = \begin{cases} DistOSPF(k,RD),\ se\ restrição\ (d)\ ou\ (e)\ for\ possível \\ DistCER_{RO,RD,RA},\ se\ restrição\ (f)\ for\ possível \end{cases}, \quad se\ (j)\ for\ possível \quad \text{(l)}$$

$$k \in ALT_{RO,RD,RA}\ ;\ RA \notin Vizinhos\_RO \quad \text{(m)}$$

Para identificar qual o $S\_CER_{RO,RD,RA}$ para um *RD*, a função custo $Z'_{RO,RD,RA}$ minimiza o valor do $custo\_CER_{RO,RD,RA}$ com o número de roteadores $num\_CER_{RO,RD,RA}$. A localização do *RC* é obtida por análise de cada roteador na seqüência de roteadores de $ALT_{RO,RD,RA}$, i.e. variável *k*. Esta variável torna-se *RC*, o que indica o último roteador do $CER_{RO,RD,RA}$ na restrição *(j)*, se uma das restrições na seqüência for aceita: *(d)* para nível ECMP, *(e)* para nível LFA ou *(f)* para nível SIG. No nível SIG, *k* não pertence ao conjunto dos roteadores vizinhos a *RO* (*Vizinhos_RO*), pois *k* sempre é um roteador após o *RV* em $ALT_{RO,RD,RA}$. Se nenhuma dessas restrições for aceita, então *k* pertence ao $CER_{RO,RD,RA}$, i.e. restrição *(i)*, e outro *k* na seqüência de $ALT_{RO,RD,RA}$ deve ser analisado. Restrição *(g)* impede a seleção de $ALT_{RO,RD,RA}$ se existir uma seqüência de roteadores no caminho OSPF de *k* até *RD* que utiliza *RA*, se *RA≠RD*. A restrição *(h)* impede a seleção de $ALT_{RO,RD,RA}$ se existir uma seqüência de enlaces no caminho OSPF de *k* até *RD* que utiliza algum enlace pertencente a um SRLG. Restrição *(k)* define um valor para $num\_CER_{RO,RD,RA}$ de acordo com o nível escolhido: 1 para o nível ECMP *(d)*, o número de roteadores no caminho OSPF, i.e. $NumRoteadoresOSPF$ de *k* até *RD*, para o nível LFA *(e),* ou o número de roteadores no caminho CER (*RV* até *RC*), i.e. $NumRoteadoresCER_{RO,RD,RA}$, para o nível SIG *(f)*. O nível ECMP atribui o valor 1 na restrição *(k)* com o objetivo de obter múltiplos caminhos de recuperação de mesmo custo para permitir balanceamento de carga com ECMP já adotado no OSPF. A restrição *(l)* também define o valor de $custo\_CER_{RO,RD,RA}$ de acordo com o nível escolhido: soma dos custos dos enlaces do caminho OSPF de *k* até *RD* (*DistOSPF*) para os níveis ECMP e LFA, ou a soma dos custos dos enlaces do caminho CER de *RV* até *RC*, i.e. $DistCER_{RO,RD,RA}$, para o nível SIG.

Portanto, a formulação obtém, para cada *RD*, o menor $S\_CER_{RO,RD,RA}$ de todos os caminhos alternativos de $\varphi_{RO,RD,RA}$ possíveis e com o menor número de roteadores.

Um algoritmo foi proposto para calcular $Z_{RO,J,RA}$ e $Z'_{RO,RD,RA}$. Obtém-se o conjunto $\varphi_{RO,RD,RA}$ e por fim identifica o $S\_CER_{RO,RD,RA}$ de forma distribuída (em cada roteador). O $S\_CER_{RO,RD,RA}$ selecionado para cada *RD* é armazenado em um vetor: *CER_Vetor*[*RD*].

O caminho obtido pelo E-CER é sempre o mesmo menor caminho obtido pelo OSPF após o término de convergência se a falha for do *RA*. Essa afirmação é provada pela geração pró-ativa dos CERs, que sempre contornam o *RA* quando possível (*RA≠RD*, pois contorna automaticamente *(RO,RA)*), mesmo se uma falha real for apenas do *(RO,RA)*. Com isso, o E-CER permite com sucesso uma adaptação gradual de todos ou de grande parte dos roteadores do CER (de *RV* ao *RC*) aos novos caminhos OSPF obtidos durante o período de convergência sem a alteração do sentido das rotas.

2.2. Extensão E-CER na FIB

Uma vez que os $S\_CER_{RO,RD,RA}$ para todos os *RD* foram gerados em cada roteador e armazenados em *CER_Vetor*[*RD*], o E-CER deve adicionar uma extensão na FIB para representar esses vetores. Essa extensão será utilizada imediatamente durante a ocorrência de falha. Todas as abordagens relacionadas também possuem uma extensão similar, mas no caso do E-CER, essa extensão consiste de um par utilizado na marcação de pacotes (*CER_Marca / Interface de Rede* (*IR*)). Esse par é utilizado por um processo de encaminhamento auxiliar ao encaminhamento IP padrão nomeado *CER_EncDif*.

2.2.1 Marcação *CER_Marca/IR*

Uma *CER_Marca* é a representação de marca para pacotes, que é definido por cada roteador (*RO*) durante o cálculo dos CERs e é relacionado a um *CER_Vetor*[*RD*] selecionado. Essa marca é representada com 16 bits divididos em 10 bits para identificar o *RO* (*RO_id*), e 6 bits para representar um *CER_Vetor*[*RD*] selecionado (*CER_id*). O *RO_id* possibilita no máximo 1024 roteadores IGP em uma mesma área OSPF, com no máximo 64 *CER_id* possíveis por roteador. A limitação de 1024 roteadores por área OSPF é mais que suficiente, pois, em situações práticas, uma área OSPF definida com mais de 1024 roteadores não é escalável devido ao alto tráfego de informações do estado do enlace. Os 6 bits para *CER_id* são suficientes, pois em nossos testes com representação de topologias artificiais e reais foram necessários no máximo 40 diferentes *CER_ids* em roteadores isolados e em média foram necessários 16 *CER_ids*.

Para representar os 16 bits *CER_Marca* no pacote IP, evitou-se utilizar mecanismos de encapsulamento tanto no cabeçalho IPv4 quanto no IPv6. Para tanto o E-CER procura reusar campos inutilizados desses cabeçalhos:

- IPv4: Os 16 bits do campo *Fragment Offset*, com a padronização da abordagem *Path MTU Discovery* (PMTU) [Mathis and Heffner 2007];
- IPv6: Usam-se os 16 últimos bits do campo *Flow Label* utilizando a definição dos três primeiros bits com "111" para uso futuro [Tang et al. 2002].

O *IR* é representado com 8 bits, possibilitando até no máximo 255 interfaces de rede por roteador, sendo suficiente para os *hardwares* de roteadores atuais.

Um par *CER_Marca/IR* poderia ser obtido para cada *CER_Vetor*[*RD*] gerado, porém uma seleção desses vetores é realizada para reduzir a quantidade de

CER_Marca/IR. Para gerar esses pares, deve-se inicialmente seguir a duas regras:

- Todas as entradas existentes na FIB (geradas pelo OSPF) com prefixos de rede anunciados por um mesmo *RD* (informação obtida da base de informações do estado do enlace) devem usar o mesmo *CER_Vetor*[*RD*]. Neste caso, todas estas entradas na FIB devem referenciar um único par *CER_Marca/IR* que representa esse vetor, pois todos os pacotes, quando desviados, deverão seguir o mesmo CER para *RD*;
- Se existir dois ou mais *CER_Vetor* com diferentes *RD*, porém com a mesma seqüência de roteadores, então todas as entradas na FIB com os prefixos anunciados por estes *RDs* devem referenciar um mesmo par *CER_Marca/IR*. Isto é possível uma vez que o CER a ser usado é o mesmo para os diferentes *RDs* necessitando apenas de uma marca;

Após observar essas duas regras, os pares *CER_Marca/IR* são gerados e adicionados na FIB seguindo um dos possíveis casos: *RO_caso* e *Não_RO_caso*.

O *RO_caso* ocorre quando um roteador gera seus próprios CERs, sendo esse roteador o *RO* no cálculo dos CERs. Nesse caso, para cada *CER_Vetor*[*RD*] selecionado uma *CER_Marca* é gerada contendo como *RO_id* os últimos 10 bits do endereço *loopback* do *RO* (*Router ID*) [Moy, 1998], o que é possível, pois os *Router Ids* são geralmente organizados em um mesmo grupo de endereços IP usado para os roteadores IGP. Se o *CER_Vetor*[*RD*] é obtido pelo nível ECMP ou LFA, então o *CER_id* contém 6 bits com valor "0". Caso contrário (nível SIG), o *CER_id* é gerado durante a execução do algoritmo para cálculo dos CERs através de um número identificador, que é incrementado a cada *CER_Vetor*[*RD*] de nível SIG selecionado. O *IR* a ser utilizado é a identificação da interface do *RO* conectada ao primeiro roteador do *CER_Vetor*[*RD*], i.e. *RV*.

O *Não_RO_caso* ocorre quando um roteador pertence à seqüência de roteadores de um *CER_Vetor*[*RD*] de nível SIG gerado por outro roteador. Neste caso, um roteador pertencente ao *Não_RO_caso* pode ser o *RV, roteadores_intermediários,* ou *RC*. Portanto, o par *CER_Marca/IR* é gerado utilizando o *CER_Vetor*[*RD*] e a *CER_Marca* obtidos a partir do processo *CER-EXT*, que é detalhado na seção 2.2.3. Com esses dados, cada roteador pertencente ao *Não_RO_caso* (*RV, roteadores_intermediários, RC*) somente necessita identificar o *IR* para gerar o par *CER_Marca/IR*. Para tanto, há de se considerar um roteador Y, como sendo um dos roteadores *Não_RO_caso*, esse roteador consegue facilmente identificar *IR* como sendo a interface de Y conectada ao próximo roteador após Y na seqüencia de roteadores de *CER_Vetor*[*RD*].

Com os pares *CER_Marca/IR* gerados, a Figura 1 ilustra como estes pares são representados na FIB. Existem três pares referenciados pelas entradas existentes na FIB, mas para simplificar o exemplo, somente duas entradas são representadas com os endereços de destino anunciados pelo mesmo roteador (*D*).

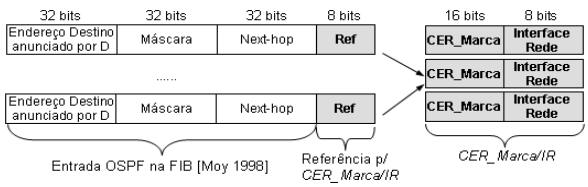

Figura 1. Representação da *CER_Marca/IR* na FIB

Estas duas entradas referenciam um mesmo par *CER_Marca/IR*, o que reduz o número necessário de marcas. A referência para esses pares é obtida através de um campo adicional *Ref* de 8 bits, o qual é suficiente para referenciar todos os *CER_Marca/IR* gerados pelo *RO_caso* e *Não_RO_caso.* Em nossos experimentos com representações de topologias artificiais e reais, apenas em alguns casos de roteadores isolados há a necessidade de no máximo de 6 bits para *Ref*, mas em média são necessários apenas 4 bits.

É importante salientar que somente os prefixos de rede anunciados pelo OSPF adicionam essa referência *Ref,* pois os demais prefixos de rede que são anunciados pelo *Border Gateway Protocol* (BGP) no interior de uma área reutilizam os prefixos definidos pelo OSPF para alcançar o *next-hop* BGP. Em caso de múltiplas áreas OSPF, o E-CER utiliza apenas a base de informações do estado do enlace de sua área, e os roteadores do tipo *Area Border Gateway* [Moy 1998] anunciam um sumário dos prefixos de rede de outras áreas. Dessa forma esses roteadores agem como *RD* na formulação sendo os anunciadores desses prefixos de rede na sua área.

Com os pares *CER_Marca/IR* definidos e preenchidos na FIB, o processo de encaminhamento *CER_EncDif* é capaz de utilizá-los na presença de falha.

2.2.2 Processo de Encaminhamento *CER_EncDif*

O processo de encaminhamento *CER_EncDif* tem por objetivo desviar os pacotes para que atinjam *RC*, e a partir desse roteador, o encaminhamento padrão baseado no endereço IP de destino é reusado pelo *CER_EncDif* para que os pacotes atinjam *RD*.

Na presença de uma falha adjacente a um roteador (torna-se o *RO*), o processo *CER_EncDif* é acionado e assume o encaminhamento para iniciar o desvio dos pacotes. O *CER_EncDif* identifica qual o *CER_Marca/IR* pela *Ref* na entrada da FIB a ser utilizada e realiza a marcação dos pacotes encaminhando-os para a respectiva *IR*. Se não existir a marca, então a topologia não possibilita 100% de cobertura de falha (seção 2).

A marcação é realizada diferentemente no IPv4 e no IPv6. No caso do IPv4, o *CER_EncDif* utiliza um bit do cabeçalho para diferenciar se o campo *Fragment Offset* está sendo usado pela marca ou pelo procedimento IPv4 de fragmentação de pacotes. O bit verificado é o primeiro bit do campo *Flags* (*Reserved Bit*), se este estiver em "0" e o campo *Fragment Offset* estiver preenchido, então um procedimento IPv4 de fragmentação foi realizado. Nesse caso, o *CER_EncDif* utiliza um mecanismo de encapsulamento para adicionar a marca, e o novo destino do pacote é o endereço do *RD* (mesmo *RD* que o endereço IP de destino original do pacote atingiria na área OSPF). Se o *Reserved Bit* estiver em "0" e o campo *Fragment Offset* estiver vazio, então o *CER_EncDif* marca o pacote com *CER_Marca,* define o *Reserved Bit* com "1" e realiza o desvio. Se o *CER_EncDif,* em um roteador ≠ *RO*, verificar que o bit está em "1" e uma falha na rota de desvio existir, isto indica que um processo de desvio já foi realizado, e nesse caso os pacotes são descartados para evitar instabilidades, pois indica a existência de múltiplas falhas independentes (diferentes do tipo SRLG, que são falhas dependentes).

No caso do IPv6, verifica-se se o campo *Flow Label* está vazio. Se estiver, o *CER_EncDif* marca o pacote definindo os três primeiros bits desse campo com "111", conforme [Tang et al. 2002], e os últimos 16 bits recebem a *CER_Marca*. Se o campo *Flow Label* está preenchido com os três primeiros bits com "111" e, em um roteador ≠

RO possuir uma falha na rota de desvio, então um processo de desvio já foi realizado com *CER_EncDif* e nesse caso, descarta-se os pacotes para evitar maiores instabilidades já que existem múltiplas falhas independentes. Caso o campo *Flow Label* esteja marcado com os três primeiros bits diferentes de "111", então houve uso do campo *Flow Label* por parte das máquinas fim a fim e nesse caso, utiliza-se o mecanismo de encapsulamento para adicionar a marca no campo *Flow Label* com os três primeiros bits com "111" e o pacote é então destinado ao *RD*.

É importante salientar que o mecanismo de encapsulamento somente é utilizado quando os campos dos cabeçalhos não puderem ser usados. Acredita-se que esses campos tendem a não ser mais usados pelos tráfegos atuais, pois, no caso do IPv4, o *Fragment Offset* é sempre evitado com o *PMTU* [Mathis and Heffner 2007]. No caso do IPv6, o *DiffServ* usa o campo *Class of Service* [Nickols et al. 1998], que melhor consegue definir parâmetros de QoS em relação ao *Flow Label*.

Se não existir uma falha adjacente, o processo *CER_EncDif* somente é acionado se um pacote marcado com *CER_Marca* entra em um roteador. Nesse caso, o *CER_EncDif* encaminha para a *IR* respectiva ao par *CER_Marca/IR*. Quando a marca do pacote não é encontrada nos pares *CER_Marca/IR* (indica que atingiu o último roteador do *CER_Vetor*, i.e. *RC*), então o encaminhamento padrão baseado apenas no endereço IP de destino (*next-hop*) é reutilizado pelo *CER_EncDif*. A marca no pacote permanece e o encaminhamento baseado no IP de destino é mantido pelo *CER_EncDif* até que atinja *RD*, onde se retira a marca dos pacotes (no caso do IPv4 define também o *Reserved Bit* com "0") ou desencapsula os pacotes. A manutenção da marca serve para o *CER_EncDif* evitar novos desvios em caso de múltiplas falhas independentes.

Em ambiente de múltiplas falhas independentes, o processo *CER_EncDif* evita o surgimento de instabilidades na rede identificando essa ocorrência e agindo com o descarte dos pacotes já marcados. De forma contrária, as abordagens NotVia e a LFA original revelam a possibilidade de um ambiente instável não agindo sobre esse problema.

2.2.3 CER-EXT

No caso dos *CER_Vetores* gerados pelos níveis ECMP e LFA, o *CER_EncDif* desvia os pacotes baseados na marca obtida até *RC=RV*. Entretanto, os *CER_Vetores* gerados pelo nível SIG definem uma seqüência de roteadores até um *RC* após o *RV*, e nesse caso o *CER_EncDif* desses roteadores necessita encaminhar os pacotes baseado na *CER_Marca/IR* para que alcancem o *RC*. Esse desvio é necessário para evitar um comportamento instável gerado pelo encaminhamento padrão IP com rotas OSPF. Para tanto, o *CER_EncDif* necessita de informações para que consiga realizar o desvio com a *CER_Marca*. Essas informações são obtidas pelo processo *CER-EXT*.

O *CER-EXT* é uma extensão do cálculo dos CERs realizado em cada roteador definido na seção 2.1. Para descrever seu funcionamento, há de se considerar um roteador X, que após realizar o cálculo dos CERs, ele deve identificar quais *CER_Vetores* de nível SIG gerados pelos demais roteadores utilizam X como *RV, roteadores_intermediários* ou *RC*. O roteador X então simula cada roteador na base de informações do estado do enlace, diferente de X, como sendo o *RO* e aplica um cálculo dos CERs para obter apenas os *CER_Vetores* que usam X na seqüência de roteadores. Em seguida, X pode identificar qual o par *CER_Marca/IR* a ser usado pelos *CER_Vetores* encontrados

através do *Não_RO_caso* (descrito na seção 2.2.1). A *CER_Marca* é, com isso, a mesma obtida pelos roteadores remotos que geram os *CER_Vetores* no *RO_caso* e também as mesmas no *Não_RO_caso* sem a necessidade de troca de mensagens de rede entre eles.

2.2.4 Manutenção do processo *CER_EncDif*

Quando uma falha é detectada, o processo *CER_EncDif* é acionado no *RO* para marcar e encaminhar os pacotes de acordo com a *IR*. O processo *CER_EncDif* mantém esse encaminhamento no *RO* por um período de tempo suficiente para que todos os processos OSPF dos roteadores consigam atualizar suas FIBs. Este período de tempo pode ser aproximadamente definido utilizando o intervalo gerado pela abordagem *oFIB* [Francois and Bonaventure 2006]. A abordagem *oFIB* define um intervalo de tempo para atualizar a FIB no roteador adjacente à falha (*RO*), sendo que esse intervalo é suficiente para os demais roteadores da topologia afetados indiretamente por essa falha (árvore SPF afetada) completem a atualização de suas FIBs. Somente após esse intervalo o *RO* atualiza a FIB e então o processo *CER_EncDif* interrompe automaticamente a marcação e desvio dos pacotes no *RO*, pois a entrada na FIB contém agora um *next-hop* correto não acionando o *CER_EncDif* para desviar os pacotes.

Somente quando o *CER_EncDif* interrompe sua marcação no *RO* que um novo cálculo dos CERs é realizado para obter *CER_Vetores* de acordo com o novo estado da topologia. Os demais roteadores da topologia, já com as rotas atualizadas na FIB, também realizam um novo cálculo dos CERs apenas quando não existir mais o fluxo de pacotes marcados e desviados em *RO*. Os pacotes marcados são encaminhados com a marca (ao atingir *RD* a marca é retirada), e se não existir a marca, então são corretamente encaminhados baseados no IP de destino (agora com as rotas atualizadas na FIB).

Durante o processo de desvio, se os pacotes marcados afetarem outros fluxos não afetados pela falha (possibilidade de congestionamentos), o *CER_EncDif* impede essa ocorrência com a seguinte regra: somente encaminhar os pacotes marcados se o tamanho da fila for menor que 80%, caso contrário, deve-se descartar esses pacotes. Em todos os testes simulados, o valor definido de 80% foi provado ser suficiente para impedir o comprometimento dos demais fluxos de pacotes quando o *CER_EncDif* estava atuando.

3. Avaliação do E-CER

Um algoritmo para E-CER foi implementado para gerar os *CER_Vetores* por cada roteador de uma topologia obedecendo à abordagem descrita na seção 2. A complexidade do algoritmo é $O(n^3 log(n))$, já incluindo o processo *CER-EXT*, onde n é o número de roteadores na área OSPF. Devido às restrições de espaço, maiores detalhes do algoritmo não são revelados. Esse algoritmo foi adaptado no simulador JavaSim (http://www.j-sim.org). A escolha desse simulador deve-se ao código OSPF portado do *GNU Zebra Project* (http://www.zebra.org), que demonstrou ser muito próximo do comportamento real do OSPF. Como o esquema E-CER utiliza as informações do estado do enlace já atualizadas pelo OSPF, o cálculo dos CERs é realizado em *background* e não influencia o processo de encaminhamento das rotas OSPF. Com os resultados desse cálculo disponíveis na FIB (*CER_Marca/IR*), se ocorrer uma falha, apenas o processo *CER_EncDif* é executado. Para realizar a avaliação, somente a abordagem NotVia foi implementada seguindo a descrição de [Bryant et al. 2007], pois essa é a

única abordagem que atinge 100% de cobertura de falha de enlace, roteador, ou SRLG.

Ao contrário do E-CER, o NotVia apenas recomenda o uso das abordagens originais de ECMP ou LFA [Bryant et al. 2007], não adaptando as mesmas para melhor utilizá-las (por exemplo, a abordagem LFA não identifica qual o melhor caminho de recuperação se existirem vários possíveis, já a abordagem E-CER, nesse caso, seleciona o caminho que conter o menor número de roteadores).

Foram geradas 50 topologias artificiais com o gerador BRITE (http://www.cs.bu.edu/brite) obedecendo aos mesmos parâmetros usados em [Hansen et al. 2006] e 8 topologias reais de *backbone* (Abilene, AGIS, ATHome, AT&T, CAIS, GEANT2, Qwest e Servint) para avaliar o E-CER em relação ao NotVia. A primeira avaliação baseia-se em uma avaliação conduzida por [Hansen et al. 2006], que confronta a extensão dos caminhos de recuperação em termos do número de roteadores percorridos entre o roteador origem até o roteador destino. Esta comparação permite identificar qual a extensão seguida pelos pacotes desviados.

A Figura 2 mostra os resultados dessa primeira avaliação para ambas as abordagens. As 50 topologias artificiais geradas com BRITE geraram resultados similares e são sintetizados em uma figura apenas. Em todas as topologias, E-CER resulta em caminhos freqüentemente menores que NotVia, pois os pacotes desviados com NotVia seguem um caminho encapsulado até *next-next-hop*, e somente a partir desse roteador, os pacotes seguem o caminho correto até o roteador destino. Esse fato tem uma maior ocorrência nas topologias AT Home, AGIS e CAIS, pois possuem seqüências de roteadores em série, o que força um caminho maior para desviar dos pacotes até o *next-next-hop* com NotVia. O E-CER, no pior caso, pode gerar o mesmo caminho que o NotVia e esse fato pode ocorrer quando existe uma alta conectividade entre os roteadores, no entanto, essa ocorrência é dependente de topologia e não foi apresentada nas topologias testadas.

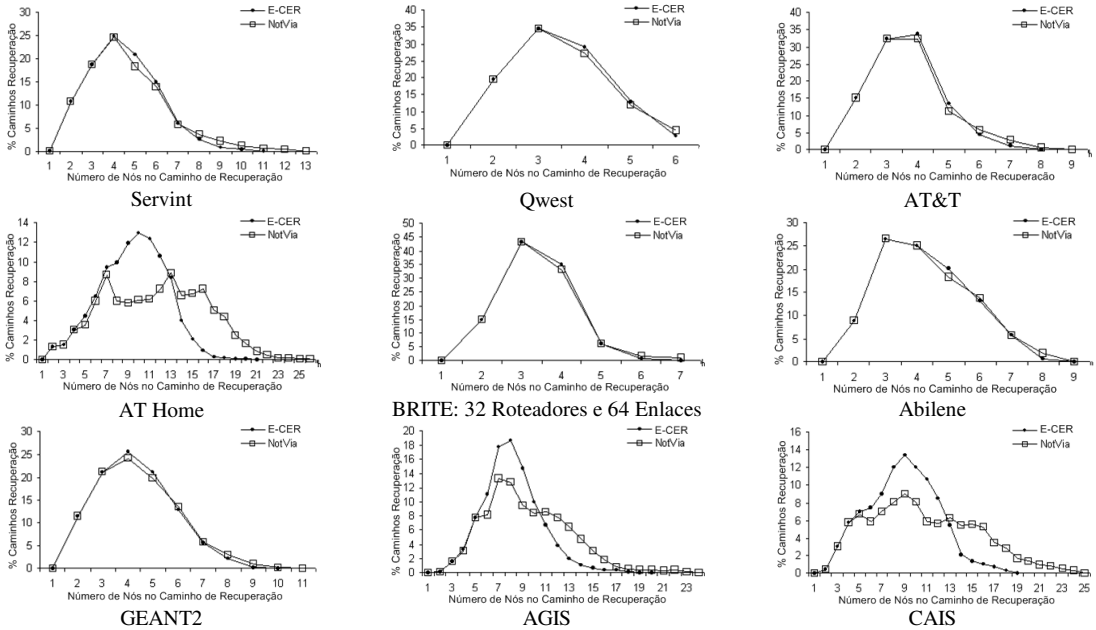

Figura 2. Número de Roteadores nos Caminhos de Recuperação por Topologia

Como a extensão do número de roteadores nos *CER_Vetores* e a geração dos

pares *CER_Marca/IR* são minimizados na seção 2, a quantidade de informações adicionadas na FIB é conseqüentemente reduzida. A segunda avaliação é a quantidade de informações adicionadas na FIB pelo E-CER e pelo NotVia. A Tabela 1 revela uma estimativa da quantidade de informações extras adicionadas na FIB por roteador, pois não foi possível obter os dados reais sobre a quantidade de entradas na FIB por roteador das topologias usadas. Usa-se nessa Tabela 1 o seguinte padrão: o número de entradas OSPF na FIB (nEO), *CER_Marca/IR* (CIR) com 3 bytes (8 bytes + 1 byte), *Ref* (seção 2.2.1) com 1 byte, cada nova entrada na FIB (nFIB) necessita de 12 bytes [Moy 1998], e um *next-next-hop* (*n-n-h*) é um endereço IP com 4 bytes [Bryant et al. 2007]. A abordagem E-CER gera uma quantidade de informações reduzida na FIB comparada com o NotVia, pois o E-CER necessita do *CER_Marca/IR* referenciado entre as entradas OSPF na FIB através do campo *Ref*, enquanto que o NotVia necessita de novas entradas na FIB para comportar o *n-n-h* além da adição desses *n-n-h* para cada entrada OSPF na FIB.

Tabela 1. Quantidade de Informação Extra adicionada na FIB

	E-CER	NotVia
Servint	12(CIR) + nEO(Ref), Total: 36 + nEO bytes	58(nFIB) + nEO(n-n-h) , Total: 696 + 4(nEO) bytes
Qwest	10(CIR) + nEO(Ref), Total: 30 + nEO bytes	93(nFIB) + nEO(n-n-h) , Total: 1116 + 4(nEO) bytes
AT&T	11(CIR) + nEO(Ref), Total: 33 + nEO bytes	72(nFIB) + nEO(n-n-h), Total: 864 + 4(nEO) bytes
AT Home	36(CIR) + nEO(Ref), Total: 108 + nEO bytes	106(nFIB) + nEO(n-n-h), Total: 1272 + 4(nEO) bytes
BRITE	10(CIR) + nEO(Ref), Total: 30 + nEO bytes	128(nFIB) +nEO(n-n-h) , Total: 1536 + 4(nEO) bytes
Abilene	9(CIR) + nEO(Ref), Total: 27 + nEO bytes	30(nFIB) + nEO(n-n-h) , Total: 360 + 4(nEO) bytes
GEANT2	12(CIR)+ nEO(Ref), Total: 36 + nEO bytes	76(nFIB) + nEO(n-n-h) , Total: 912 + 4(nEO) bytes
AGIS	31(CIR) + nEO(Ref), Total: 93 + nEO bytes	149(nFIB) +nEO(n-n-h) , Total: 1788 + 4(nEO) bytes
CAIS	30(CIR) + nEO(Ref), Total: 90 + nEO bytes	88(nFIB) +nEO(n-n-h) , Total: 1056 + 4(nEO) bytes

A terceira avaliação verifica qual a taxa de redução de pacotes perdidos durante o período de convergência do OSPF quando com o uso da abordagem E-CER. Algumas mudanças foram necessárias no JavaSim para modificar o código OSPF de forma a ter um período de convergência aproximado de 200ms além de retardar o processo de sinalização de falha em 20ms, conforme [Francois et al. 2005] (antes de 20ms, os pacotes são descartados). Utilizou-se apenas a representação da topologia GEANT2 (ilustrado na Figura 3) para essa análise, pois se conhecia as capacidades reais dos seus enlaces (gigabit), de forma que os seus pesos são valores inversamente proporcionais às suas capacidades e o período de convergência do OSPF pode alcançar 200 ms [Francois et al. 2005].

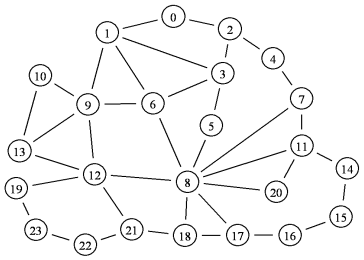

Figura 3. GEANT2

Planejou-se uma matriz de tráfego de forma que todos os enlaces da topologia ocupassem no máximo 50% de suas capacidades (ver seção 2). Entre os fluxos de tráfego, a avaliação ocorre em 6 pares de fluxos (origem, destino): (2,18), (18,2), (1,17), (17,1), (9,11) e (11,9), pois estes usam dois dos roteadores centrais da topologia (6 e 8)

sendo o pior ambiente de falha. Esses 6 fluxos são gerados a uma taxa de bits constante de 160Mbps e pacotes com tamanho 256 bytes para ajustar o enlace 6-8 a 480 Mbps.

Definiram-se 3 cenários de falha: enlace 6-8, roteador 6, e roteador 8. Primeiramente a avaliação ocorre com o OSPF (convergência ≈ 200ms) e depois com E-CER auxiliando o OSPF durante esse período de convergência. Todos os fluxos de pacotes analisados são desviados e conseguem se ajustar à largura de banda remanescente (50% livre). Dessa forma, é possível analisar isoladamente a taxa de pacotes perdidos gerado exclusivamente durante o período de convergência, o que não pode ser medido corretamente em um ambiente instável com uma alta taxa de tráfego na topologia. Entretanto, o *CER_EncDif* consegue impedir um ambiente mais instável (com congestionamentos) descartando os pacotes marcados no momento que a capacidade das filas for > 80% (seção 2.2.4). Uma regra similar não existe em qualquer outra abordagem IPFRR, porém é uma medida recomendada [Shand and Bryant 2007].

A Tabela 2 revela a porcentagem de pacotes perdidos durante o período de convergência. O OSPF tem a maior taxa de pacotes perdidos devido à sua característica reativa, mesmo com um período de convergência de ≈ 200ms. A abordagem E-CER auxilia o OSPF durante esse período, pois durante ≈ 200ms, o processo *CER_EncDif* realiza o encaminhamento para desviar os pacotes baseados no par *CER_Marca/IR*, o que explica a baixa taxa de pacotes perdidos. Após 200ms, os processos OSPF dos roteadores adjacentes à falha são os últimos a atualizar as rotas na FIB (utilizando *oFIB*, ver seção 2.2.4) e o *CER_EncDif* não é mais acionado.

Tabela 2. % Pacotes Perdidos Durante o Período de Convergência

	OSPF (≈ 200ms)			E-CER + OSPF (≈ 200ms)		
	Enlace 6-8	Roteador 6	Roteador 8	Enlace 6-8	Roteador 6	Roteador 8
Fluxo (2,18)	23,4 %	24,5 %	25,4 %	1,8 %	2,8 %	2,7 %
Fluxo (18,2)	23,2 %	24,1 %	25 %	1,4 %	2,6 %	2,6 %
Fluxo (1,17)	23,5 %	24,7 %	24,7 %	1,6 %	2,7 %	2,5 %
Fluxo (17,1)	23,1 %	24,7 %	25,1 %	1,5 %	2,6 %	2,6 %
Fluxo (9,11)	23,1 %	24,3 %	25,3 %	2,1 %	2,5 %	2,7 %
Fluxo (11,9)	23,2 %	24,2 %	25,4 %	1,3 %	2,6 %	2,6 %

4. Conclusão

Dependendo da aplicação utilizada, os protocolos de roteamento não reagem à uma falha em tempo hábil, mesmo reduzindo o período de convergência em menos de 1 segundo. Isto pode acarretar em uma alta taxa de pacotes perdidos comprometendo a confiabilidade da rede. Propõe-se a nova abordagem E-CER para gerar previamente caminhos de recuperação como auxílio ao OSPF. O E-CER toma como base o *framework* IPFRR, porém realizando adaptações para minimizar os caminhos e a quantidade de roteadores necessários para desvio dos pacotes. Essa característica resulta em caminhos de recuperação menores, com reduzida representação na FIB e que se mostrou capaz de reduzir a taxa de pacotes perdidos durante o período de convergência. Dependendo da falha, cada caminho obtido com o E-CER é sempre o mesmo caminho obtido com OSPF após o período de convergência, o que permite uma adaptação gradual dos roteadores às novas rotas do OSPF. Uma implementação em ambiente real, bem como aspectos de segurança e uma versão para tratar múltiplas falhas independentes serão abordados em trabalhos futuros.

References

Ahuja, R. K., Magnanti, T. L., Orlin, B. O. (1993) Network Flows: Theory, Algorithms, and Applications, Prentice-Hall.

Atlas, A., and Zinin, A. (2007). "Basic Specification for IP Fast-Reroute Loop-Free Alternate". In *Internet Draft*. IETF Routing-WG.

Barreto, F., Wille, E. C. G., and Nacamura Jr, L. (2007) "Contornando Falhas em Backbones IP com Caminhos Emergenciais Rápidos", In *SBRC 2007 Workshop de Teste e Tolerância a Falhas*.

Basu, A. and Riecke, J. F. (2001). "Stability Issues in OSPF Routing". In *SIGGCOM Applications, technologies, arch., and protocols for communications*, pages 225-236.

Bryant, S., FilsFils, S., Previdi, S., and Shand, M. (2005). "IP Fast Reroute Using Tunnels". In *Internet Draft*. IETF Routing-WG.

Bryant, S., Shand, M., and Previdi, S. (2007). "IP Fast Reroute Using Not-via Address". In *Internet Draft*. IETF Routing-WG.

Francois, P., and Bonaventure, O. (2006). "Avoiding Transient Loops during IGP Convergence in IP Networks". In *IEEE INFOCOM Computer Communications*, pp. 237-247

Francois, P., Filfis, C., Evans, C., and Bonaventure, O. (2005). "Achieving sub-second IGP convergence in large IP networks". In *ACM SIGCOMM*, pages 34-44.

Hansen, A. F., and Cicic, T., and Gjessing, S. (2006). "Alternative Schemes for Proactive IP Recovery". In *Next Generation Internet Desing and Engineering,* pages 1-8.

Iannaccone, G., Chuah, C., Bhattacharyya, S., and Diot, C. (2004). "Feasibility of IP Restoration in a Tier-1 Backbone". In *IEEE Network Magazine*, pages 13-19.

Kvalbein, A., Hansen, A. F., Cicic, T., Gjessing, S., and Lysne, O. (2006). "Fast IP Network Recovery using Multiple Routing Configurations". In *IEEE INFOCOM Computer Communications*, pages 1-11.

Mathis, M., and Heffner, J. (2007). "Packetization Layer Path MTU Discovery". In *RFC 4821*, IETF Network-WG.

Moy, J. (1998). "OSPF version 2". In *RFC 2328*. IETF Network-WG.

Narvaez, P. (2000). Routing Reconfiguration in IP Networks. Phd Thesis. Massachusetts Institute of Technology.

Nelakuditi, S., Lee, S., Yu, Y., and Zang, Z. (2007) "Fast Local Rerouting for Handling Transient Link Failures". In *IEEE Transactions on Networking*, pages 359-372.

Nickols, K., Blake, S., Baker, F. Black, D. (1998) "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers". In *RFC 2474*. IETF Network-WG.

Shand, M., and Bryant, S. (2007). "IP Fast Reroute Framework". In *Internet Draft*. IETF Routing-WG.

Tang, X., Tang, J., Huang, G., and Siew, C. (2002). "QoS Provisioning Using IPv6 Flow Label in the Internet". In *IEEE Conference on Communication and Signal Processing*.