

4 TRABALHOS RELACIONADOS

Neste capítulo serão descritas de forma sucinta três propostas de tolerância à falha dos Agentes de Mobilidade. As duas primeiras propostas empregam redundâncias de hardware para alcançar este objetivo. Já a terceira utiliza um mecanismo mais complexo, em virtude de não propor a utilização de redundância de hardware. Como a nova proposta é baseada em melhorias realizadas nesta terceira proposta serão citadas somente as diferenças existentes entre estas duas.

Em síntese o artigo da referência bibliográfica [16] propõe um mecanismo de tolerância à falha dos Agentes de Mobilidade empregando redundâncias de hardware, ou seja, cada Agente de Mobilidade possui múltiplos agentes como membros de backup. Cada um desses agentes monitora constantemente um ao outro e quando alguma falha ocorre em um deles, um outro agente o substitui fazendo o seu trabalho. Para evitar o desperdício de uma máquina (membro de backup) operar somente após a ocorrência da falha da máquina principal, também é proposto que esses membros operem simultaneamente com a máquina principal. Logo a carga de tráfego deste domínio de rede será dividida entre essas máquinas (load-sharing model).

Outra proposta feita neste artigo é que seja distribuída, entre os membros de backup, as listas que associam os endereços IP das diferentes estações móveis de um mesmo domínio de rede com os seus respectivos Care-of-addresses (CoAs). Isto é feito da seguinte forma: as mensagens de Solicitação e Resposta de Registro que são trocadas entre as estações móveis e os seus respectivos Agentes de Origem são estendidas para seus membros de backup. Isto é, as estações móveis registram-se somente com um Agente de Origem, mas os demais agentes membros de backup deste HA “ouvem” essas mensagens. A “distribuição” dessas listas de associações de endereços IP para os membros de backup desse HA é feita com o intuito de evitar que uma estação móvel tenha que registrar-se com todos os membros de backup a cada novo processo de registro. Ou seja, evita que aumente o número de mensagens de Solicitação e de Resposta de Registro na rede.

Agora será explicado como este protocolo de tolerância à falha dos Agentes de Mobilidade funciona. Para isto, será exemplificado esse funcionamento com o Agente de Origem, embora os procedimentos desse funcionamento sejam os mesmos empregados na tolerância à falha dos Agentes Estrangeiros. Acompanhe a explicação observando a figura 11 abaixo.

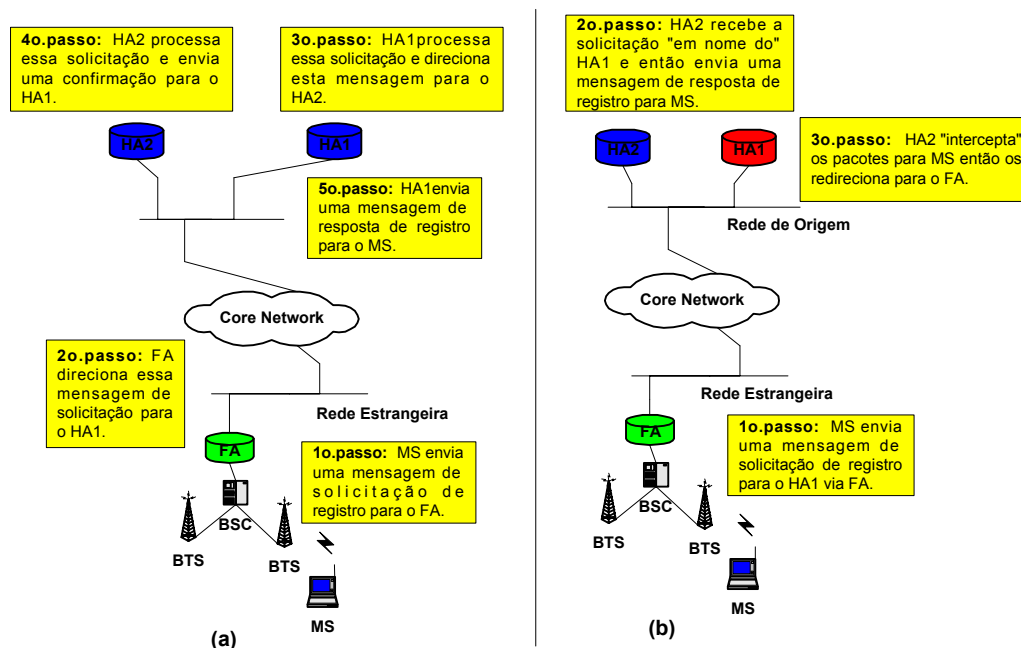


Figura 11 - Protocolo de tolerância à falha dos HAs usando redundância de hardware.

A figura 11(a) ilustra a estação móvel (MS) enviando uma mensagem de Solicitação de Registro para o Agente Estrangeiro (FA) da área de cobertura em que esta estação móvel se encontra. Essa mensagem é redirecionada para o Agente de Origem dessa estação móvel, o Agente de Origem de número 1 (HA1). Este agente por sua vez processa essa mensagem e caso este agente aceite essa solicitação de registro ele deve redirecionar essa mensagem de solicitação para o HA2 (membro de backup). O HA2 armazena em sua lista de associações de endereços IP (Mobility Binding) o endereço IP do Agente Estrangeiro (FA) e envia uma confirmação para o HA1, indicando que já processou essa informação. Após o recebimento dessa confirmação o HA1 envia uma mensagem de Resposta de Registro para a estação móvel concedendo o registro. Caso o HA1 não conceda essa solicitação de registro, este agente gera uma mensagem de Resposta de Registro negando a solicitação e o HA2 não recebe nenhum tipo de mensagem do HA1.

Quando o HA1 falha, acompanhe a figura 11(b), o HA2 passa a receber a solicitação de registro “em nome do HA1”. A estação móvel (MS) sente uma pequena interrupção no fornecimento do serviço, mas essa interrupção é bem menor do que seria caso esta estação móvel tivesse que fazer uma nova solicitação de registro, agora com o HA2. Quando o HA1 volta a funcionar este agente troca informações com o HA2 com intuito de recuperar a lista Mobility Binding.

Se existem mais de dois Agentes de Origem como membro de backup é preciso utilizar um mecanismo de decisão para determinar qual desses agentes substituirá o agente defeituoso. Para que isso seja feito é proposto neste artigo que o Agente de Origem principal e seus membros de backup fiquem constantemente monitorando um ao outro para saber, qual deles tem a menor quantidade de endereços IP. O Agente de Origem (membro de backup) que possui a menor quantidade será escolhido como novo Agente de Origem principal.

O modo de operação de carga de tráfego compartilhada (load-sharing) é conseguido configurando as diferentes estações móveis de um mesmo domínio de rede a terem como Agente de Origem principal uma das redundâncias de hardware deste domínio.

A segunda proposta de tolerância à falha que pode ser vista com mais detalhe na referência bibliográfica [17] propõe um protocolo de tolerância à falha dos Agentes Estrangeiros que constituem a rede hierárquica que tem um Gateway Foreign Agent (GFA) como o dispositivo que ocupa a maior posição hierárquica dessa rede estrangeira. O objetivo desta proposição é manter a arquitetura hierárquica dessa rede intacta na ocorrência de falhas dos Agentes Estrangeiros. Isso é conseguido posicionando redundâncias de hardware ao longo dessa rede hierárquica. Para compreendermos melhor o funcionamento deste protocolo acompanhe a explicação com base na figura 12. Nesta figura a letra Rx (onde x é um número) indica que o dispositivo é uma redundância de hardware do nível hierárquico que este dispositivo de rede se encontra. O membro de backup do FA2 é o R1, pois esse dispositivo redundante está conectado aos três Agentes Estrangeiros do nível hierárquico imediatamente abaixo. Seguindo esta mesma idéia o membro de backup do FA4 é o R2. Este artigo propõe que o GFA que nesta figura é o FA1, associe para cada FA dos níveis hierárquicos inferiores o seu correspondente membro de backup.

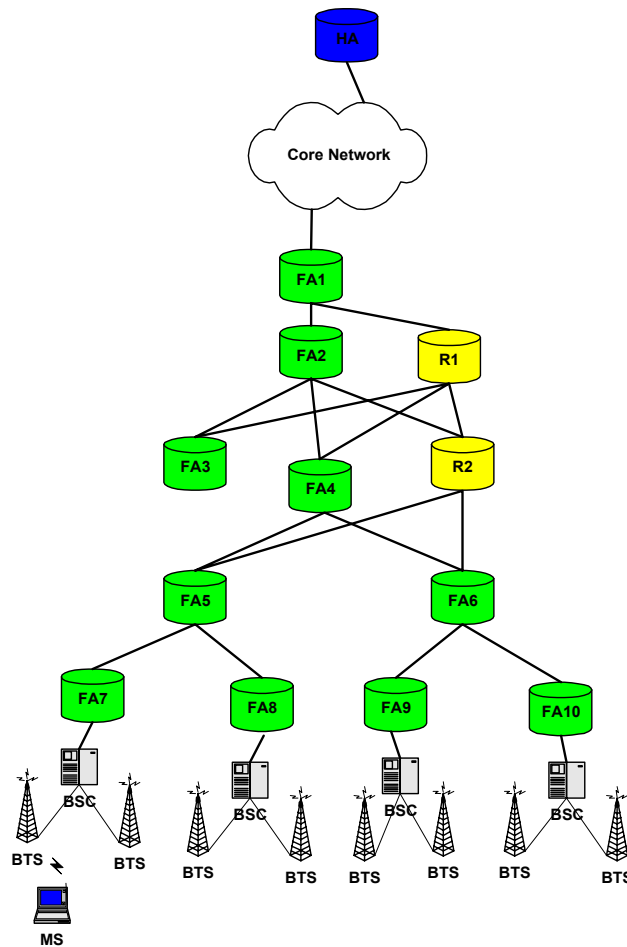


Figura 12 - Agentes Estrangeiros em uma topologia de rede hierárquica.

No cenário acima foi suposto que o FA4 parou de funcionar. Então será executado um conjunto de procedimentos para manter a arquitetura dessa rede hierárquica intacta. Esses procedimentos estão descritos nos itens logo abaixo:

- a) O FA2 que está no nível hierárquico imediatamente acima do FA4 detecta a ocorrência dessa falha.
- b) Em seguida o FA2 comunica-se com FA1 com intuito de obter a informação de qual dispositivo de rede é o membro de backup do FA4.
- c) Após o FA2 ser informado que o R2 é o membro de backup do FA4, o FA2 envia uma mensagem para o R2 com objetivo de ativá-lo para que ele desempenhe a mesma função que o FA4 desempenhava antes da falha. Através do R2, o FA2 informa aos Agentes Estrangeiros de número 5 e 6 que o FA4 falhou e que o R2 a partir deste momento irá desempenhar as funções do FA4.

- d) Em seguida o FA5 e FA6 atualizam suas rotas de roteamento para o nível hierárquico acima, ou seja, são ativadas as rotas que conectam o FA5 ao R2 e o FA6 ao R2. Após as ativações dessas novas rotas, o FA5 e o FA6 notificam o R2 que eles estão operando, isto é fornecendo conectividade às estações móveis que estão sob as áreas de cobertura das Estações Bases (BSs) conectadas aos Agentes Estrangeiros dos níveis hierárquicos inferiores. Portanto, para que o R2 saiba quais são essas estações móveis, a lista de estações móveis visitantes do FA5 e do FA6 são repassadas para o R2.
- e) Após o R2 ter obtido essas listas de estações móveis visitantes dos Agentes Estrangeiros de número 5 e 6, este membro de backup (R2) deve repassar essas listas para o FA2, pois provavelmente essas listas sofreram algumas alterações com os deslocamentos dessas estações móveis.
- f) Quando o FA2 recebe a lista de estações móveis visitantes do R2, este FA atualiza suas listas de visitantes.
- g) Em seguida o FA1 deve ser informado da mudança que ocorreu na estrutura dessa rede hierárquica, isto é, o FA2 deve informar ao FA1 que o membro de backup do R2 é agora o FA4.
- h) Como periodicamente o FA2 monitora o FA4, então após o reparo do Agente de Origem defeituoso (FA4) o FA2 detecta esse reparo. E para que essa rede volte a operar na sua estrutura hierárquica original após o reparo do FA4, o FA2 age como se tivesse ocorrido uma falha no R2, então executa todos os passos novamente, desde dos itens b) ao g).

A terceira proposta [18] de tolerância à falha dos Agentes de Mobilidade é mais complexa em virtude de utilizar uma gerência de rede que determinará qual dos Agentes de Mobilidade livre de falha serão os membros de backup dos agentes defeituosos. Essa proposta é interessante, pois os seus autores, Jenn-Wei & Joseph Arul, conseguem provar, por simulações, que quando as cargas de tráfegos dos Agentes de Mobilidade começam a aumentar chega um determinado ponto desse aumento que é mais interessante se ter membros de backup que são alocados dinamicamente na rede utilizando os Agentes de Mobilidade de outros

domínios de redes ao invés de se ter hardware que operem no modo load-sharing ou standby. Isso pode ser explicado pelo fato de que nessa proposta as cargas de tráfegos dos agentes defeituosos são distribuídas para os demais agentes livres de falhas enquanto que nas propostas que sugerem a redundância de hardware essas cargas de tráfegos estão restritas ao gerenciamento do hardware redundante. Outro ponto que merece ser destacado na comparação dessas duas propostas é o fato de que, nas propostas que sugerem o emprego de hardware redundante, se esses membros de backup falharem simultaneamente com o Agente de Mobilidade ou falharem antes da recuperação desse hardware principal, todas as estações móveis atingidas pela falha desse hardware ficarão sem serviços.

Como a nova proposta apresentada nesta dissertação é baseada em melhorias feitas na terceira proposta, serão citadas somente as diferenças existentes entre elas. A terceira proposta propõe como parâmetro para a determinação dos membros de backup a carga de tráfego disponível nos Agentes de Mobilidade. Isto é, a carga de tráfego dos Agentes Estrangeiros defeituosos (na tolerância à falha dos FAs) e das estações móveis atingidas pelas falhas dos seus respectivos Agentes de Origem (na tolerância à falha dos HAs) serão distribuídas para os agentes de mesmo tipo (HA ou FA) que possuem as menores cargas de tráfego nos momentos em que essas falhas ocorrem. Já a nova proposta emprega dois parâmetros para determinação de tais membros de backup, a carga de tráfego disponível nos Agentes de Mobilidade livres de falhas e o retardo introduzido aos pacotes de dados ao terem esses agentes como possíveis membros de backup.

Na nova proposta de tolerância à falha dos Agentes Estrangeiros é proposta a utilização de um retardo limiar que servirá de parâmetro de comparação. Este parâmetro verificará se os retardos introduzidos aos pacotes de dados que partem ou destinam-se as Estações Base Transceptoras (BTSs) atingidas pela falha de um FA ao terem outros FAs (livres de falhas) como membros de backup são aceitáveis para os serviços da classe conversacional de tempo real, onde as exigências de retardos são mais rigorosas. Portanto, essa proposta de tolerância à falha tende a determinar para cada BTS atingida pela falha do seu FA a opção de membro de backup que possui a maior carga de tráfego disponível (FA de menor sobrecarga em seu buffer) e que introduz aos pacotes de dados um valor de retardo aceitável (menor ou igual ao retardo limiar).

Na proposta de tolerância à falha dos Agentes de Origem, a principal diferença existente entre a nova proposta e a proposta apresentada por Jenn-Wei e Joseph Arul [18] é que, nessa segunda, os usuários afetados pela falha de seu Agente de Origem (HA) tentarão ser alocados no HA de maior carga de tráfego disponível (HA com menor sobrecarga o seu buffer). Porém nem sempre esta opção de membro de backup é a melhor em termos de retardos introduzidos aos pacotes de dados. Pois para algumas BTSs essa opção de membro de backup pode introduzir um retardo maior aos pacotes de dados ao serem comparadas com outras opções de Agentes de Origem ativos existentes. Portanto, na nova proposta para cada BTS serão investigadas todas as opções de Agentes de Origem ativos com intuito de determinar qual é a melhor opção quanto aos retardos introduzidos aos pacotes de dados.

E em seguida as estações móveis afetadas pela falha do seu HA e que se encontram sob a área de cobertura dessa BTS serão analisadas quanto à introdução de suas cargas de tráfegos neste HA de melhor opção quanto ao retardo médio introduzido aos pacotes de dados. Como os usuários das próximas gerações dos sistemas celulares passarão a consumir com maior intensidade diferentes tipos de serviços de dados, logo serão registradas várias taxas de dados como carga de tráfego. Porém, para que a gerência de rede responsável pelos procedimentos de determinação dos membros de backup do HA defeituoso saiba que classes de serviços àquelas estações móveis afetadas pela falha de seu HA estavam usufruindo no momento da ocorrência dessa falha, é proposto a introdução de um campo de informação que tenha esse propósito tanto na lista que associa os endereços IP das diferentes estações móveis com os seus respectivos Care-of-Addresses (lista existente nos HAs - Mobility Binding) como na lista de estações móveis visitantes em uma rede estrangeira (lista existente nos FAs).

Assim, toda vez que as listas dos FAs forem atualizadas também serão atualizadas as informações de quais tipos de serviços os diferentes usuários estão consumindo. Logo, no momento em que a gerência de rede consegue recuperar essas listas de associações dos endereços IP das estações móveis com os seus respectivos CoAs, também é recuperada a informação de qual tipo de serviço cada usuário afetado pela falha de seu HA estava utilizando no momento da ocorrência da falha. Assim, no momento em que são executados os procedimentos para alocações das estações móveis atingidas pela falha do seu HA serão asseguradas

que as estações que usufruem dos serviços da classe conversacional não fiquem expostas a um retardo maior do que as estações que usufruem dos serviços da classe interativa.

Além da proposição desse novo campo de informação nas duas listas citadas acima, é proposto que a gerência de rede recupere a lista Mobility Binding e que esta gerência seja responsável pela divisão do conteúdo dessa lista entre os HAs determinados como membro de backup por essa mesma gerência. Na proposta de Wei e Arul essa recuperação e essa divisão são feitas por um Agente de Origem escolhido pela gerência de rede. O intuito dessa nova proposição é deixar que todos os procedimentos para as determinações dos membros de backup dos Agentes de Mobilidade sejam feitos pela gerência de rede.