Introdução ao IGRP

Contents

Introduction

Este documento introduz o Interior Gateway Routing Protocol (IGRP). Ele tem dois propósitos. Um é para formar uma introdução para a tecnologia IGRP, para aquelas que estão interessados em usar, avaliar e, possivelmente, implementá-lo. A outro é para dar uma exposição mais ampla a algumas ideias e conceitos interessantes que são personificados no IGRP. Consulte Configuração do IGRP, A Implementação do Cisco IGRP e Comandos do IGRP para obter informações sobre como configurar o IGRP.

Objetivos para IGRP

O protocolo IGRP permite que vários gateways coordenem o roteamento. Seus objetivos são os seguintes:

- Roteamento estável até mesmo em redes muito grandes ou complexas. Não deve haver loops de roteamento, mesmo como transientes.
- Resposta rápida às alterações na topologia da rede.

- Overhead baixo. Ou seja, o próprio IGRP não deve usar mais largura de banda que o limite realmente necessário para sua tarefa.
- Divisão de tráfego entre várias rotas paralelas quando elas são, em termos gerais, equivalentes ao desejado.
- Considerar as taxas de erros e o nível de tráfego em caminhos diferentes.

A implementação atual do IGRP processa o roteamento para TCP/IP. No entanto, o projeto básico é destinado a lidar com uma variedade de protocolos.

Uma única ferramenta não vai resolver todos os problemas de roteamento. Convencionalmente, o problema de roteamento está dividido em várias partes. Protocolos, como o IGRP, são chamados de "protocolos de gateway interno" (IGPs). Eles foram planejados para uso em um único conjunto de redes, seja em um único gerenciamento ou em gerenciamentos coordenados. Esses conjuntos de redes estão conectados por "Protocolos de gateways externos" (EGPs). Um IGP foi projetado para supervisionar a riqueza de detalhes da topologia de rede. A prioridade do projeto de um IGP é produzir rotas ideais e responder rapidamente às mudanças. Um EGP tem a intenção de proteger um sistema de redes contra erros ou interpretação incorreta intencional por outros sistemas; o BGP é um protocolo de gateway externo desse tipo. A prioridade em projetar um EGP está na estabilidade e nos controles administrativos. Freqüentemente, é suficiente que um EGP produza uma rota razoável, em vez da rota ideal.

O IGRP possui algumas similaridades com protocolos mais antigos, como o Routing Information Protocol da Xerox, RIP da Berkeley e o Hello de Dave Mills. Ele se distingue desses protocolos principalmente por ser projetado para redes maiores e mais complexas. Consulte a seção Comparação com RIP para obter uma comparação mais detalhada com RIP, que é o mais amplamente utilizado dos protocolos da geração mais antiga.

Como esses protocolos mais antigos, o IGRP é um protocolo de vetor de distância. No protocolo, os gateways trocam informações de roteamento apenas com gateways adjacentes. Essas informações de roteamento contêm um resumo das informações sobre o resto da rede. Pode ser comprovado matematicamente que todos os gateways juntos solucionam um problema de otimização, de acordo com um certo montante para um algoritmo distribuído. Cada gateway precisa apenas resolver parte do problema e apenas receber uma parte do total de dados.

A principal alternativa ao IGRP é o EIGRP (IGRP Avançado) e uma classe de algoritmos chamada SPF (caminho mais curto primeiro). O OSPF usa esse conceito. Para saber mais sobre OSPF, consulte Guia de design do OSPF. O OSPF se baseia em uma técnica de inundação, onde cada gateway sempre é mantido atualizado sobre o status de cada interface em todos os outros gateways. Cada gateway soluciona de forma independente o problema de otimização do seu ponto de vista, usando dados de toda a rede. Existem vantagens em cada abordagem. Em algumas circunstâncias, o SPF pode conseguir responder às alterações mais rapidamente. Para evitar circuitos de roteamento, o IGRP precisa ignorar novos dados por alguns minutos após certos tipos de mudança. Como o SPF recebe informações diretamente de cada gateway, ele consegue evitar esses Routing Loops. Dessa forma, pode atuar imediatamente sobre as novas informações. Entretanto, o SPF precisa lidar substancialmente com mais dados que o IGRP, tanto em estruturas de dados internos quanto em mensagens entre gateways.

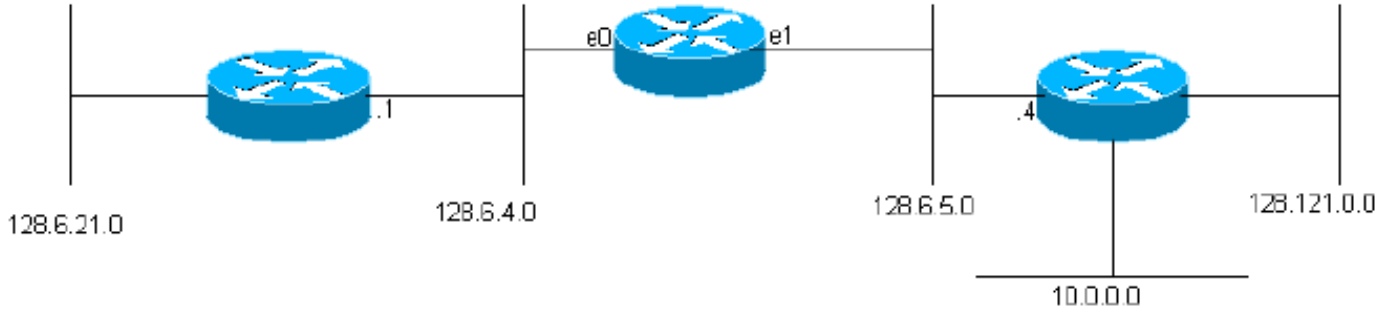
O Problema de Roteamento

O IGRP tem como objetivo ser usado em gateways conectando várias redes. Vamos supor que as redes usem a tecnologia baseada em pacotes. Na realidade, os gateways atuam como switches de pacotes. Quando um sistema conectado a uma rede precisar enviar um pacote para

um sistema em uma rede diferente, ele endereçará o pacote para o gateway. Se o destino estiver em uma das redes conectadas ao gateway, este encaminhará o pacote ao destino. Se o destino for mais distante, o gateway encaminhará o pacote para outro gateway mais próximo ao destino. Os gateways usam tabelas de roteamento para ajudá-los a decidir o que fazer com os pacotes. Aqui está uma tabela de roteamento simples como exemplo. (Os endereços usados nos exemplos são endereços IP obtidos da Rutgers University. Observe que o problema básico de roteamento é semelhante em outros produtos também, mas essa descrição pressupõe que o IGRP esteja sendo utilizado para IP Routing.)

Figure 1

```
network      gateway     interface
  -------      -------     ---------
  128.6.4      none        ethernet 0
  128.6.5      none        ethernet 1
  128.6.21     128.6.4.1   ethernet 0
  128.121      128.6.5.4   ethernet 1
  10           128.6.5.4   ethernet 1
```



(As tabelas de roteamento IGRP reais têm informações adicionais para cada gateway, conforme mostrado a seguir.) Esse gateway está conectado a duas redes Ethernet, chamadas 0 e 1. Eles receberam números de rede de IP (na realidade, números de sub-rede) 128.6.4 e 128.6.5. Assim, os pacotes endereçados para essas redes específicas podem ser enviados diretamente para o destino, usando a interface Ethernet apropriada. Há dois gateways próximos, o 128.6.4.1 e o 128.6.5.4. Os pacotes para redes diferentes de 128.6.4 e 128.6.5 serão encaminhados para um desses gateways ou para outros. A tabela de roteamento indica qual gateway deve ser usado e para qual rede. Por exemplo, pacotes endereçados a um host na rede 10 devem ser encaminhados para o gateway 128.6.5.4. Espera-se que esse gateway esteja mais próximo à rede 10, isto é, que o melhor caminho para a rede 10 passe por este gateway. O principal objetivo do IGRP é permitir que os gateways criem e mantenham tabelas de roteamento como essa.

Resumo de IGRP

Conforme mencionado acima, o IGRP é um protocolo que permite aos gateways construírem a tabela de roteamento trocando informações com outros gateways. Um gateway é iniciado com entradas para todas as redes conectadas diretamente a ele. Ele obtém informações sobre outras redes trocando atualizações de roteamento com gateways adjacentes. No caso mais simples, o gateway encontrará um caminho que representa a melhor maneira de chegar a cada rede. Um caminho é caracterizado pelo próximo gateway ao qual os pacotes devem ser enviados, a interface de rede que deve ser utilizada e informações de métricas. As informações de métricas são um conjunto de números que caracterizam o nível de viabilidade do caminho. Isso permite que o gateway compare os caminhos obtidos de vários gateways e decida o caminho a ser utilizado. Geralmente, há situações em que faz sentido dividir o tráfego entre dois ou mais

caminhos. IGRP fará isso sempre que dois ou mais caminhos forem igualmente bons. O usuário também poderá configurá-lo para dividir o tráfego quando os caminhos forem igualmente viáveis. Nesse caso, mais tráfego será enviado junto com o caminho com a melhor métrica. A intenção é que o tráfego possa ser dividido entre uma linha de 9600 bps e outra de 19200 bps e a linha 19200 obterá aproximadamente duas vezes mais tráfego que a linha de 9600 bps.

As métricas usadas pelo IGRP incluem o seguinte:

- Atraso relativo à topologia
- A largura de banda do segmento com a menor largura de banda do caminho
- Ocupação do canal do caminho
- Confiabilidade do caminho

O tempo de atraso topológico é a quantidade de tempo que levaria para chegar até o destino ao longo desse caminho, supondo que uma rede esteja sem carga. Há, obviamente, retardo adicional quando a rede é carregada. No entanto, a carga é contabilizada usando o número de ocupação do canal, e não tentando medir atrasos reais. A largura de banda do caminho é simplesmente a largura de banda por segundo do link mais lento do caminho. A ocupação de canal indica a quantidade de largura de banda em uso no momento. Ela é medida e será alterada com carga. A confiabilidade indica a taxa de erros atual. Ela é a fração de pacotes que chegam ao destino sem danos. Ela é medida.

Embora não sejam usadas como parte da métrica, duas partes de adição de informações são transmitidas com ela: contagem e MTU de salto. A contagem de salto é apenas o número de gateways pelo qual um pacote terá que passar para chegar ao destino. O MTU é o tamanho de pacote máximo que pode ser enviado ao longo de todo o caminho, sem fragmentação. (Ou seja, é o mínimo de MTUs de todas as redes envolvidas no caminho.)

Com base nas informações métricas, uma única "métrica composta" é calculada para o caminho. A métrica composta combina o efeito de vários componentes de métrica em um único número que representa a "viabilidade" desse caminho. Ela é a métrica composta que, na verdade, é usada para decidir qual é o melhor caminho.

Periodicamente, cada gateway transmite sua tabela de roteamento inteira (com uma certa censura devido à regra de horizonte de divisão) a todos os gateways adjacentes. Quando um gateway obtém a transmissão de outro gateway, ele compara a tabela com sua tabela atual. Todos os destinos novos e caminhos são adicionados à tabela de roteamento do gateway. Os caminhos na transmissão são comparados aos caminhos atuais. Se um novo caminho for melhor, ele poderá substituir o atual. As informações na difusão também são utilizadas para atualizar a ocupação do canal e outras informações sobre caminhos existentes. Esse procedimento geral é semelhante ao utilizado por todos os protocolos de vetor de distância. Ele é conhecido nas especificações da literatura matemática como o algoritmo de Bellman-Ford. Consulte o RFC 1058 para obter um desenvolvimento detalhado do procedimento básico, que descreve o RIP, um protocolo de vetor de distância mais antigo.

No IGRP, o algoritmo geral de Bellman-Ford é modificado em três aspectos críticos. Primeiro, em vez de uma métrica simples, um vetor de métricas é usado para caracterizar os caminhos. Segundo, em lugar de escolher um único caminho com a menor métrica, o tráfego é dividido entre vários caminhos, cujas métricas caem em uma faixa especificada. Terceiro, vários recursos são apresentados para fornecer estabilidade nas situações em que a topologia está mudando.

O melhor caminho é selecionado com base em uma métrica composta:

```
[(K1 / Be) + (K2 * Dc)] r
```
Em que K1, K2 = constantes, Be = largura de banda do pacote descarregado x (1 ocupação do canal), Dc = retardo topológico e r = confiabilidade.

O caminho que tiver a métrica composta menor será o melhor caminho. Onde há vários caminhos para o mesmo destino, o gateway pode rotear os pacotes por mais de um caminho. Isso é feito de acordo com a métrica composta para cada caminho de dados. Por exemplo, se um caminho tiver uma métrica composta de 1 e outro caminho tiver uma métrica composta de 3, três vezes mais pacotes serão enviados pelo caminho de dados que tenha a métrica composta de 1.

Há duas vantagens em relação ao uso de um vetor de informações de métrica. A primeira é que ele dá a capacidade de suportar vários tipos de serviço do mesmo conjunto de dados. A segunda vantagem é a precisão aprimorada. Quando uma métrica única é usada, ela normalmente é tratada como se houvesse um atraso. Cada link no caminho é adicionado à métrica total. Se houver um link com uma largura de banda baixa, ele normalmente é representado por um grande atraso. No entanto, as limitações da largura de banda não se acumulam da mesma maneira que os atrasos. Ao tratar a largura de banda como um componente separado, ela pode ser usada corretamente. De maneira semelhante, a carga pode ser controlada por um número de ocupância de canal separado.

O IGRP fornece um sistema de interconexão de redes de computadores que pode lidar de modo estável com uma topologia de gráfico gerais, incluindo loops. O sistema mantém informações de métrica de todo caminho, ou seja, ele sabe os parâmetros do caminho para todas as outras redes sem gateway conectado. O tráfego pode ser distribuído ao longo de caminhos paralelos, e parâmetros de caminho múltiplo podem ser computados simultaneamente em toda a rede.

Comparação com RIP

Esta seção compara IGRP com RIP. Essa comparação é útil porque o RIP é amplamente utilizado para finalidades similares ao IGRP. Entretanto, fazer isso não é totalmente justo. O RIP não foi destinado a atender a todos os mesmos objetivos de IGRP. O RIP foi criado para uso em redes pequenas com a tecnologia razoavelmente uniforme. Nesses aplicativos, geralmente é adequado.

A diferença mais básica entre IGRP e RIP é a estrutura das suas métricas. Infelizmente, esta não é uma mudança que possa ser simplesmente encaixada retroativamente no RIP. Ela exige os novos algoritmos e as estruturas de dados presentes em IGRP.

O RIP utiliza uma métrica simples de "contagem de nós" para descrever a rede. Ao contrário do IGRP, onde cada caminho é descrito por um atraso, largura de banda etc., no RIP, ele é descrito por um número de 1 a 15. Normalmente esse número é usado para representar quantos gateways são percorridos pelo caminho, antes de obter o destino. Isso significa que não há nenhuma distinção entre uma linha serial lenta e uma Ethernet. Em algumas implementações do RIP, o administrador do sistema pode especificar que um determinado salto deve ser contado mais de uma vez. As redes lentas podem ser representadas por uma grande contagem de salto. Mas, como o máximo é 15, isso não pode ser feito com muita frequência. Por exemplo, se uma Ethernet é representada por 1 e uma linha de 56 Kb por 3, pode haver no máximo 5 linhas de 56 Kb em um caminho, senão o número máximo de 15 será excedido. Para representar toda a variedade de velocidades de rede disponível e permitir uma rede grande, estudos feitos pela Cisco sugerem que é necessária uma métrica de 24 bits. Se a métrica máxima for pequena demais, o administrador do sistema terá duas opções desagradáveis: ou ele não pode distinguir

entre rotas rápidas e lentas, ou ele não consegue encaixar sua rede completa no limite. Na realidade, várias redes nacionais agora são grandes o suficiente para serem tratadas pelo RIP, mesmo se cada salto for contado somente uma vez. O RIP simplesmente não pode ser utilizado para estas redes.

A resposta óbvia seria modificar o RIP para permitir uma métrica maior. Infelizmente, isso não vai funcionar. Como todos os protocolos de vetor de distância, o RIP tem o problema de "contar até o infinito". Isso é descrito em mais detalhes no RFC 1058 . Quando a topologia muda, rotas artificiais são introduzidas. As métricas associadas a essas rotas artificiais lentamente aumentam até que atinjam 15, e, nesse ponto, as rotas são removidas. 15 é um valor máximo pequeno suficiente para que esse processo seja convergido razoavelmente rápido, supondo que as atualizações acionadas sejam usadas. Se o RIP foi modificado para permitir uma métrica de 24 bits, os loops persistiriam tempo suficiente para que a métrica seja contada até 2**24. Isso não é tolerável. O IGRP tem recursos projetados para impedir que rotas artificiais sejam introduzidas. Isso é discutido abaixo na seção 5.2. Não é prático lidar com redes complexas sem introduzir esses recursos, ou alterar para um protocolo, como o SPF.

O IGRP faz bem mais do que simplesmente aumentar o intervalo de métricas permitidas. Reestrutura a métrica para descrever o retardo, a largura de banda, a confiabilidade e a carga. É possível representar essas considerações em uma única métrica, por exemplo, RIPs. No entanto, a abordagem aplicada pelo IGRP é potencialmente mais precisa. Por exemplo, com uma única métrica, vários links rápidos e sucessivos parecerão equivalentes a um único link lento. Esse pode ser o caso do tráfego interativo, em que o atraso é a principal preocupação. No entanto, para transferências de dados em grande escala, a preocupação principal é a largura de banda e a adição de métricas em conjunto não é o método certo aqui. O IGRP processa o atraso e a largura de banda separadamente, acumulando atrasos, mas utilizando o mínimo das larguras de banda. Não é fácil saber como incorporar os efeitos de confiabilidade e carga em uma métrica de componente único.

Na minha opinião, uma das grandes vantagens do IGRP é a facilidade de configuração. Ele pode representar diretamente as quantidades que tenham significado físico. Isso significa que ela pode ser configurada automaticamente, com base no tipo de interface, na velocidade da linha, etc. Com uma métrica de componente único, é mais provável que a métrica tenha que ser "cozinhada" para incorporar efeitos de várias coisas diferentes.

Outras inovações se referem mais a algoritmos e estruturas de dados do que ao Routing Protocol. Por exemplo, o IGRP especifica algoritmos e estruturas de dados que suportam divisão de tráfego entre várias rotas. Certamente, é possível projetar uma implementação do RIP que faça isso. Entretanto, uma vez o roteamento tendo sido reimplementado, não há motivo para manter o RIP.

Até agora eu descrevi o "IGRP genérico", uma tecnologia que poderia suportar o roteamento para qualquer protocolo de rede. Entretanto, nesta seção é válido falar um pouco mais sobre a implementação específica de TCP/IP. Essa é a implementação que será comparada ao RIP.

As mensagens de atualização de RIP contêm apenas instantâneos da tabela de roteamento. Ou seja, elas têm vários destinos e valores métricos. A implementação de IP do IGRP tem uma estrutura adicional. Primeiro, a mensagem de atualização é identificada por um "número de sistema autônomo". Essa terminologia provém da tradição Arpanet e tem aqui um significado específico. Entretanto, em muitas redes isto significa que é possível executar vários sistemas de roteamento diferentes na mesma rede. Isso é útil para locais em que há convergência de redes de várias empresas. Cada organização pode manter seu próprio roteamento. Como cada atualização é rotulada, os gateways podem ser configurados para observar apenas a correta. Certos gateways são configurados para receber atualizações de diversos sistemas autônomos.

Eles passam informações entre os sistemas de maneira controlada. Observe que esta não é uma solução completa para os problemas de segurança do roteamento. Qualquer gateway pode ser configurado para ouvir atualizações de qualquer sistema autônomo. No entanto, essa continua sendo uma ferramenta muito útil na implementação de políticas de roteamento nas quais há um grau razoável de confiança entre os administradores de rede.

O segundo recurso estrutural sobre mensagens de atualização de IGRP afeta a maneira como as rotas padrão são tratadas pelo IGRP. A maioria dos protocolos de roteamento possui um conceito de rota padrão. Geralmente, não é prático que as atualizações de roteamento listem todas as redes do mundo. Geralmente, um conjunto de gateways precisa de informações detalhadas de roteamento para as redes da organização. Todo o tráfego para destinos fora de sua empresa pode ser enviado para um dos poucos gateways limítrofes. É possível que esses gateways limite tenham informações mais completas. A rota para o melhor gateway limítrofe é uma “rota padrão”. É um padrão no sentido em que é usado para chegar a qualquer destino que não esteja listado especificamente nas atualizações de roteamento internas. O RIP, e alguns outros protocolos de roteamento, circulam informações sobre a rota padrão como se ela fosse uma rede real. O IGRP usa um método diferente. Em vez de uma única entrada falsa para a rota padrão, o IGRP permite que redes reais sejam marcadas como candidatas para ser um padrão. Isto é implementado pelo posicionamento de informações sobre essas redes em uma seção externa especial da mensagem de atualização. No entanto, também pode ser considerado como o ativador de um bit associado a essas redes. Periodicamente, o IGRP faz a varredura de todas as rotas padrão candidatas e escolhe aquela com a menor métrica para ser a rota padrão atual.

Potencialmente, essa aproximação dos padrões é um tanto mais flexível que a aproximação usada pela maioria das implementações de RIP. A maioria dos gateways tipicamente RIP pode ser definida para gerar uma rota padrão com uma certa métrica especificada. A intenção é fazer isso nos gateways de limite.

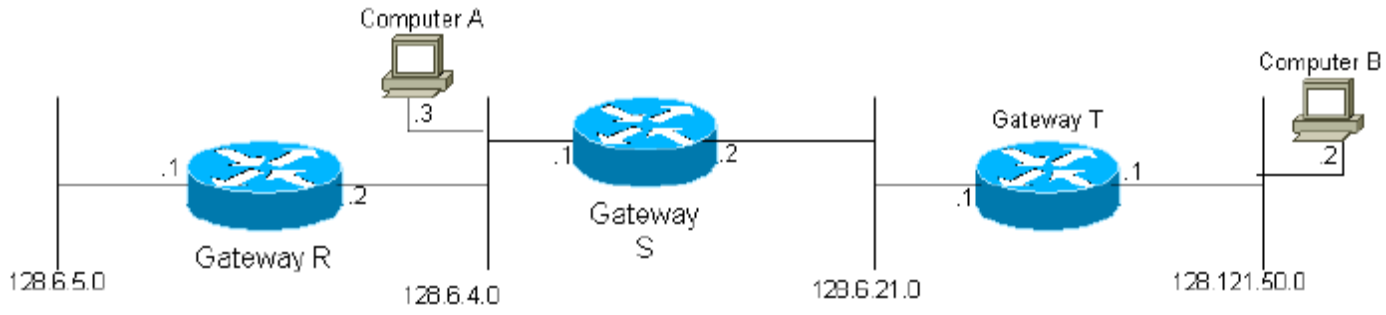
Descrição detalhada

Esta seção fornece uma descrição detalhada do IGRP.

Descrição geral

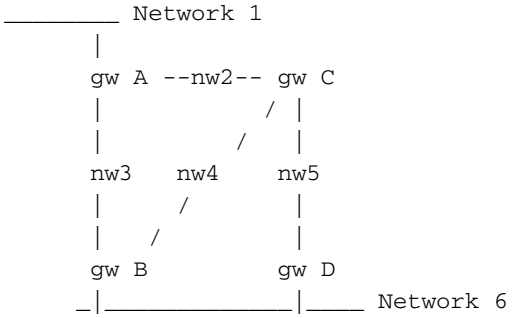
Quando um gateway é ligado pela primeira vez, a tabela de roteamento respectiva é inicializada. Isso pode ser feito por um operador a partir do terminal do console ou lendo as informações nos arquivos de configuração. Uma descrição de cada rede conectada ao gateway é fornecida, incluindo o atraso topológico junto ao link (por exemplo, quanto tempo demora para um único bit atravessar o link) e a largura de banda do link.

Figure 2



Por exemplo, no diagrama acima, o gateway S seria informado de que está conectado às redes 2 e 3 por meio das interfaces correspondentes. Portanto, inicialmente, o gateway 2 só sabe que pode acessar qualquer computador de destino nas redes 2 e 3. Todos os gateways são programados para transmitir periodicamente aos seus gateways vizinhos as informações com as quais foram inicializados, bem como as informações coletadas de outros gateways. Assim, o gateway S receberia atualizações dos gateways R e T e aprenderia que ele pode acessar os computadores na rede 1 através do gateway R e os computadores na rede 4 através do gateway T. Como o gateway S envia toda a tabela de roteamento, no próximo ciclo, o gateway T aprenderá que ele pode chegar à rede 1 através do gateway S. É fácil perceber que as informações sobre todas as redes no sistema acabarão alcançando todos os gateways no sistema, desde que a rede esteja completamente conectada.

Figure 3



Cada gateway calcula uma métrica composta para determinar a conveniência dos caminhos de dados para computadores de destino. Por exemplo, no diagrama acima, para um destino na Rede 6, o gateway A (gw A) calcularia as funções de métrica para dois caminhos, através dos gateways B e C. Observe que os caminhos são definidos simplesmente pelo próximo salto. Na realidade, há três rotas possíveis do A para a Rede 6:

- Direto para B
- Para C e, depois, para B
- Para C e, em seguida, para D

No entanto, o gateway A não precisa escolher entre as duas rotas que envolvem C. A tabela de roteamento em A tem uma única entrada que representa o caminho para C. Sua métrica representa a melhor maneira de chegar de C ao destino final. Se A envia um pacote para C, fica a critério de C decidir utilizar B ou D.

Equação 1

A função métrica composta calculada para cada caminho de dados é mostrada abaixo:

```
[(K1 / Be) + (K2 * Dc)] r
```

Onde r = confiabilidade fracionária (% de transmissões que são recebidas com sucesso no próximo salto), Dc = atraso composto, Be = largura de banda efetiva: largura de banda sem carga x (1 - ocupação de canal) e K1 e K2 = constantes.

Equação 2

Em princípio, o atraso composto, Dc, poderia ter sido determinado como mostrado abaixo:

```
Dc = Ds + Dcir + Dt
```

Onde Ds = retardo de switching, Dcir = retardo do circuito (retardo da propagação de 1 bit) e DT = retardo da transmissão (nenhum retardo de carregamento para uma mensagem de 1500 bits)

No entanto, na prática, um número de atraso padrão é usado para cada tipo de tecnologia de rede. Por exemplo, haverá uma figura de atraso padrão para Ethernet e para linhas seriais em qualquer taxa de bits específica.

Eis um exemplo de como pode ser a tabela de roteamento do gateway A no caso do diagrama de rede 6 acima. (Observe que componentes isolados do vetor de métrica não são mostrados, para efeito de simplicidade.)

Exemplo de tabela de roteamento:

Rede	Interface	Próximo gateway	Métrico
1	NW 1	Nenhum	Diretamente conectado
2	NW 2	Nenhum	Diretamente conectado
3	NW3	Nenhum	Diretamente conectado
4	NW 2	C	1270
	NW3	B	1180
5	NW 2	C	1270
	NW3	B	2130
6	NW 2	C	2040
	NW3	B	1180

O processo básico de criação de uma tabela de roteamento trocando informações com os vizinhos é descrito pelo algoritmo Bellman-Ford. O algoritmo foi usado em protocolos anteriores, como o RIP (RFC 1058). Para lidar com redes mais complexas, o IGRP adiciona três recursos ao algoritmo básico de Bellman-Ford:

1. No lugar de uma métrica simples, um vetor de métrica é usado para caracterizar os caminhos. Uma única métrica composta pode ser calculada com base nesse vetor, de acordo com a Equação 1, acima. O uso de um vetor permite que o gateway acomode diferentes tipos de serviço, usando vários coeficientes diferentes na Equação 1. Também permite uma representação mais precisa das características da rede do que uma métrica única.
2. Em vez de escolher um caminho simples com a menor métrica, ele divide o tráfego em vários caminhos cujas métricas se encontram em um intervalo especificado. Isso permite que várias rotas sejam usadas em paralelo, fornecendo uma largura de banda efetiva maior do que qualquer rota única. Uma variação V é especificada pelo administrador da rede. Todos os caminhos com métrica composta mínima M são mantidos. Além disso, todos os caminhos cuja métrica for inferior a V x M são mantidos. O tráfego é distribuído entre vários caminhos em proporção inversa às métricas compostas.
3. Há alguns problemas com esse conceito de variância. É difícil encontrar estratégias que

façam uso de valores de variância maiores que 1 e que também não levem a looping de pacotes. No Cisco versão 8.2, o recurso de variância não está implementado. (Não tenho certeza em qual versão o recurso foi removido.) O efeito disso é definir a variância permanentemente como 1.

4. Diversos recursos são introduzidos para oferecer estabilidade em situações nas quais a topologia está se modificando. Esses recursos destinam-se a evitar loops de roteamento e "contar até infinito", que caracterizaram as tentativas anteriores de usar algoritmos do tipo Ford para esse tipo de aplicativo. Os recursos de estabilidade primária são “holddowns”, “triggered updates” (atualizações disparadas), “split horizon” e “poisoning”. Eles serão discutidos mais detalhadamente abaixo.

A divisão do tráfego (ponto 2) aumenta um perigo sutil. A variação V é designada para permitir que os gateways utilizem caminhos paralelos de velocidades diferentes. Por exemplo, é possível que haja uma linha de 9600 BPS sendo executada em paralelo com uma linha de 19200 BPS, para redundância. Se a variância V for 1, somente o melhor caminho será usado. Portanto, a linha 9600 BPS não será usada se a linha 19200 BPS tiver uma confiabilidade razoável. (No entanto, se vários caminhos forem iguais, a carga será compartilhada entre eles.) Aumentando a variância, podemos permitir que o tráfego seja dividido entre a melhor rota e outras rotas quase tão boas quanto. Com uma grande variação, o tráfego será dividido entre as duas linhas. O perigo é que com uma variação de tamanho suficiente, os caminhos que se tornem alocados não sejam exatamente os mais lentos, mas que estejam realmente, na direção errada. Portanto, deveria haver uma regra adicional para evitar que o tráfego fosse enviado upstream: Nenhum tráfego é enviado aos caminhos cuja métrica composta remota (a métrica composta calculada no próximo salto) é superior à métrica composta calculada no gateway. Em geral, os administradores do sistema são encorajados a não definir a variação acima de 1, exceto em situações específicas em que os caminhos paralelos precisem ser usados. Nesse caso, a variação é cuidadosamente configurada para fornecer os resultados certos.

O IGRP foi planejado para cuidar de vários tipos de serviço e vários protocolos. O tipo de serviço é uma especificação em um pacote de dados que modifica a maneira de avaliar os caminhos. Por exemplo, o protocolo TCP/IP permite que o pacote especifique a importância relativa da alta largura de banda, do fraco atraso ou da alta confiabilidade. Geralmente, as aplicações interativas especificarão um retardo baixo, enquanto as aplicações de transferência de grande escala especificarão uma alta largura de banda. Esses requisitos determinam os valores relativos de K1 e K2 que são apropriados para uso em Eq. 1. Cada combinação de especificações no pacote que deve ser suportado é mencionada como um tipo de serviço. Para cada tipo de serviço deve ser escolhido um conjunto de parâmetros K1 e K2. Uma tabela de roteamento é mantida para cada tipo de serviço. Isso é feito porque os caminhos são selecionados e ordenados de acordo com a métrica composta definida por Eq. 1. Isso é diferente para cada tipo de serviço. As informações de todas essas tabelas de roteamento são combinadas para produzir mensagens atualizadas de roteamento intercambiadas pelos gateways, como descrito na Figura 7.

Recursos de estabilidade

Esta seção descreve holddowns, atualizações disparadas, split horizon e envenenamento. Esses recursos são projetados para impedir que os gateways selecionam rotas erradas. Conforme descrito no RFC 1058 , isso pode acontecer quando uma rota se torna inutilizável, devido à falha de um gateway ou de uma rede. A princípio, os gateways adjacentes detectam falhas. Então, eles enviam atualizações de roteamento que exibem a rota antiga como não-utilizável. No entanto, é possível que as atualizações não cheguem a algumas partes da rede ou demorem a chegar em certos gateways. Um gateway que ainda considera a rota antiga como adequada pode continuar a disseminar essas informações, reinserindo assim a rota defeituosa no sistema. Eventualmente,

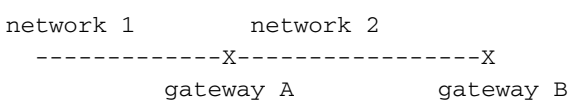
essas informações se propagam pela rede e retornam ao gateway que as injetou novamente. O resultado é uma rota circular.

Na verdade, há alguma redundância entre as contramedidas. Em princípio, holddowns e atualizações disparadas devem ser suficientes para evitar rotas errôneas no primeiro lugar. Na prática, entretanto, falhas de comunicação de diversos tipos pode fazer com que sejam insuficientes. O horizonte dividido e o envenenamento de rota destinam-se a evitar loops de roteamento em qualquer caso.

Normalmente, novas tabelas de roteamento são regularmente enviadas para gateways vizinhos (a cada 90 segundos por padrão, embora isso possa ser ajustado pelo administrador do sistema). Uma atualização iniciada é uma nova tabela de roteamento que é enviada imediatamente, em resposta a alguma mudança. A alteração mais importante é a remoção de uma rota. Isso pode acontecer porque um tempo limite se esgotou (provavelmente um gateway ou linha vizinha ficaram inativos) ou porque uma mensagem de atualização do próximo gateway no caminho mostra que o caminho não é mais utilizável. Quando um gateway G detecta que uma rota não é mais aproveitável, ele dispara uma atualização imediatamente. Essa atualização mostrará que a rota não é utilizável. Considere o que acontece quando essa atualização atinge os gateways vizinhos. Se a rota do vizinho apontar de volta para G, o vizinho deve remover a rota. Isso faz com que o vizinho dispare uma atualização, etc. Assim, uma falha acionará uma onda de mensagens de atualização. Essa onda se propagará por toda a parte da rede, em que as rotas passaram pelo gateway, ou pela rede com falha.

Atualizações disparadas seriam insuficientes se pudéssemos garantir que a onda de atualizações atingiria cada gateway apropriado imediatamente. No entanto, há dois problemas. Primeiro, os pacotes que contêm a mensagem de atualização podem ser descartados ou danificados por algum link na rede. Em segundo lugar, as atualizações acionadas não acontecem instantaneamente. É possível que um gateway que ainda não tenha recebido a atualização disparada emita uma atualização normal exatamente no momento errado, fazendo com que a rota errada seja reinserida no vizinho que já recebeu a atualização disparada. Holddowns são projetados para solucionar esses problemas. A regra de holddown afirma que, quando uma rota é removida, nenhuma nova rota será aceita para o mesmo destino pelo mesmo período. Isso oferece às atualizações acionadas tempo para obter todos os outros gateways, para que possamos nos certificar de que as rotas que obtemos não são apenas algum gateway reinserindo o antigo. O período de suspensão deve ser longo o suficiente para permitir que a onda de atualizações acionada percorra toda a rede. Além disso, deve incluir alguns ciclos de transmissão comuns, para manejar pacotes perdidos. Leve em consideração o que acontece se uma das atualizações acionadas for descartada ou danificada. O gateway que emitiu essa atualização emitirá outra na próxima atualização regular. Isso reiniciará a onda de atualizações disparadas em vizinhos que perderam a onda inicial.

A combinação de atualizações acionadas e suspensões deve ser suficiente para se livrar das rotas expiradas e impedir que elas sejam reinseridas. No entanto, de qualquer maneira, vale a pena tomar algumas precauções adicionais. Elas permitem redes com muitas perdas e redes particionadas. As precauções adicionais para IGRP são split horizon e route poisoning. O Split horizon surge da observação de que nunca faz sentido enviar uma rota de volta na direção da qual veio. Considere a seguinte situação:



O gateway A dirá ao B que possui uma rota para a rede 1. Quando B envia atualizações para A,

não há motivos para mencionar a rede 1. Como A está mais perto de 1, não há razão para considerar passar por B. A regra “split horizon” diz que uma mensagem de atualização separada deve ser gerada para cada vizinho (na verdade, cada rede vizinha). A atualização de um determinado vizinho deve omitir as rotas que apontam para esse vizinho. Essa regra evita loops entre gateways adjacentes. Por exemplo, suponha que a interface de A para a rede 1 falhe. Sem a regra “split horizon”, B diria para A que pode chegar a 1. Já que não tem mais uma rota real, A pode pegar essa rota. Nesse caso, A e B teriam rotas para 1. Mas A apontaria para B e B apontaria para A. É claro que as atualizações acionadas e os holddowns devem impedir que isso aconteça. Mas, como não há motivo para retornar informações ao local de origem, o horizonte de divisão vale a pena mesmo assim. Além de sua função de impedir loops, o horizonte dividido controla o tamanho das mensagens de atualização.

O horizonte dividido deve evitar loops entre os gateways adjacentes. O envenenamento de rota destina-se a quebrar loops maiores. A regra é que, quando uma atualização mostra que a métrica de uma rota existente aumentou o suficiente, existe um loop. A rota deverá ser removida e colocada em holddown. Atualmente, a regra é a remoção de uma rota, se a métrica composta aumentar mais que um fator de 1,1. Não é seguro, para qualquer aumento na métrica composta, acionar a remoção da rota, pois pequenas alterações na métrica podem ocorrer devido a alterações na ocupação ou confiabilidade do canal. Portanto, o fator de 1.1 é somente um heurístico. O valor exato não é obrigatório. Esperamos que esta regra seja necessária apenas para quebrar loops muito grandes, já que os pequenos serão impedidos por atualizações acionadas e suspensões.

Desativar holddowns

A partir da versão 8.2, o código da Cisco fornece uma opção para desabilitar holddowns. A desvantagem dos holddowns é que eles atrasam a adoção de uma nova rota, quando uma rota antiga falha. Com os parâmetros padrão, pode levar muitos minutos para um roteador adotar uma nova rota após uma alteração. Entretanto, pelos motivos explicados acima, não é seguro simplesmente remover holddowns. O resultado seria contar até o infinito, conforme descrito em RFC 1058. Supomos, mas não podemos provar, que com uma versão mais forte do envenenamento por rota, as suspensões não são mais necessárias para interromper a contagem até o infinito. Desta maneira, desabilitar holddowns habilita essa forma mais forte de corrupção de rota. Observe que o split horizon e atualizações disparadas ainda estão vigorando.

A forma mais forte de envenenamento de rota se baseia em uma contagem de nó. Se a contagem de saltos para um caminho aumentar, a rota será removida. Obviamente, essa ação removerá as rotas que ainda são válidas. Se um elemento em algum outro lugar da rede sofrer alteração fazendo com que o caminho passe por mais um gateway, a contagem de saltos aumentará. Nesse caso, a rota ainda é válida. Entretanto, não há um meio totalmente seguro de distinguir este caso dos loops de roteamento (contagem até o infinito). Dessa maneira, a abordagem mais segura é remover a rota sempre que a contagem de saltos aumentar. Se a rota é legítima, ela será reinstalada pela próxima atualização, e isso causará uma atualização que reinstalará a rota em qualquer outro lugar do sistema.

Em geral, os algoritmos de vetor de distância 1 adotam facilmente novas rotas. O problema é limpar completamente os antigos do sistema. Assim, uma regra excessivamente agressiva sobre como remover rotas suspeitas deve ser segura.

Detalhes do processo de atualização

O conjunto de processos descrito nas Figuras 4 a 8 visa a manipular um protocolo de rede

simples, por exemplo, o protocolo TCP/IP, DECnet ou ISSO/OSI. No entanto, os detalhes do protocolo serão dados apenas para TCP/IP. Um único gateway pode processar dados que seguem mais de um protocolo. Como cada protocolo tem diferentes estruturas de endereçamento e formatos de pacotes, o código de computador usado para implementar as Figuras 4 a 8 geralmente será diferente para cada protocolo. O processo descrito na Figura 4 varia mais, conforme descrito nas notas detalhadas da Figura 4. Os processos descritos nas Figuras 5 a 8 terão a mesma estrutura geral. A principal diferença de protocolo para protocolo será o formato do pacote de atualização do roteamento, que deve ser definido como sendo compatível com um protocolo específico.

Observe que a definição de um destino pode variar de protocolo para protocolo. O método descrito aqui pode ser usado para o roteamento para hosts individuais, para redes ou para esquemas mais complexos de endereço hierárquico. O tipo de roteamento usado dependerá da estrutura de endereçamento do protocolo. A implementação atual de TCP/IP suporta apenas roteamento para redes IP. Assim, o "destino" realmente significa o número de rede ou sub-rede de IP. As informações de sub-rede são mantidas somente para redes conectadas.

As figuras 4 a 7 mostram pseudocódigos para várias partes do processo de roteamento usados pelos gateways. No início do programa, protocolos e parâmetros aceitáveis descrevendo cada interface são inseridos.

O gateway só lidará com certos protocolos listados. Qualquer comunicação de um sistema usando um protocolo que não esteja na lista será ignorada. As entradas de dados são as seguintes:

- Redes às quais o gateway está conectado.
- Largura de banda não carregada de cada rede.
- Atraso topológico de cada rede.
- Confiabilidade de cada rede.
- Ocupação de canal de cada rede.
- MTU de cada rede.

A função métrica para cada caminho de dados é calculada de acordo com a Equação 1. Observe que os três primeiros itens são razoavelmente permanentes. Eles são uma função da tecnologia de rede subjacente e não dependem da carga. Eles podem ser definidos a partir de um arquivo de configuração ou por uma entrada direta do operador. Observe que o IGRP não utiliza atraso medido. Tanto a teoria quanto a experiência sugerem que é muito difícil para protocolos que utilizam o retardo medido manter o roteamento estável. Há dois parâmetros medidos: ocupação de canal e confiabilidade. A confiabilidade se baseia nas taxas de erros reportadas pelo hardware ou firmware da interface de rede.

Além dessas entradas, o algoritmo de roteamento requer um valor para diversos parâmetros de roteamento. Isso inclui os valores do cronômetro, a variação, e se as suspensões estão ativadas. Isso normalmente seria especificado por um arquivo de configuração ou entrada de operador. (Desde o lançamento do Cisco 8.2, a variação está permanentemente definida como 1).

Depois que as informações iniciais são inseridas, as operações no gateway são acionadas por eventos - a chegada de um pacote de dados em uma das interfaces de rede ou a expiração de um cronômetro. Os processos descritos nas Figuras 4 a 7 são acionados da seguinte forma:

- Quando um pacote chega, ele é processado de acordo com a Figura 4. Isso faz com que o pacote seja enviado para outra interface, descartado ou aceito para processamento posterior.
- Quando um pacote é aceito pelo gateway para processamento adicional, ele é analisado de

uma maneira específica ao protocolo, não descrita nesta especificação. Se o pacote é uma atualização de roteamento, ele é processado de acordo com a Figura 5.
- A Figura 6 mostra eventos acionados por um cronômetro. O cronômetro é definido de forma a gerar uma interrupção uma vez por segundo. Quando a interrupção ocorre, o processo mostrado na Figura 6 é executado.
- A Figura 7 mostra uma sub-rotina de atualização de roteamento. As chamadas para esta sub-rotina são mostradas nas Figuras 5 e 6.
- Além disso, a figura 8 mostra detalhes de computação métrica mencionados nas figuras 5 e 7.

Existem quatro constantes de tempo essenciais que controlam a propagação e a expiração da rota. Essas constantes de tempo podem ser definidas pelo administrador do sistema. No entanto, há valores padrão. Estas constantes de tempo são:

- Horário da transmissão — As atualizações são transmitidas por todos os gateways em todas as interfaces conectadas com frequência. O padrão é a cada 90 segundos.
- Tempo inválido — Se nenhuma atualização foi recebida para um determinado caminho dentro desse período, considera-se que o tempo limite se esgotou. Ele deve ser várias vezes o tempo de transmissão, a fim de permitir a possibilidade de que pacotes contendo uma atualização possam ser descartados pela rede. O padrão é 3 vezes o tempo de transmissão.
- Tempo de espera — Quando um destino se torna inacessível (ou a métrica aumenta o suficiente para causar envenenamento), o destino entra em "suspensão". Nesse estado, nenhum novo caminho será aceito para o mesmo destino nesse período. O tempo de espera indica a duração desse estado. O tempo de espera deve ser várias vezes o tempo de difusão. O valor padrão é 3 vezes o tempo de broadcast mais 10 segundos. (Conforme descrito na seção Desabilitar Holddowns, é possível desabilitar holddowns.)
- Liberar tempo — Se nenhuma atualização tiver sido recebida para um determinado destino dentro desse período, a entrada para ela será removida da tabela de roteamento. Observe a diferença entre o tempo inválido e o tempo de limpeza: Depois do tempo inválido, um caminho tem o tempo esgotado e é removido. Se não houver mais caminhos para um destino, ele agora será inalcançável. No entanto, a entrada do banco de dados para o destino permanece. Ela deve permanecer para reforçar o holddown. Após o tempo de descarga, a entrada do banco de dados é removida da tabela. Ele deve ser um pouco mais longo do que o tempo de holddown. O padrão é 7 vezes o tempo de transmissão.

Estes números pressupõem as seguintes estruturas de dados principais. Um conjunto separado dessas estruturas de dados é mantido para cada protocolo suportado pelo gateway. Dentro de cada protocolo, é mantido um conjunto separado de estruturas de dados para cada tipo de serviço a ser suportado.

Para cada destino conhecido no sistema, há uma lista (nula possivelmente) de caminhos para o destino, um tempo de expiração de holddown e um tempo recente de atualização. A hora da última atualização indica a última hora em que qualquer caminho para este destino foi incluído em uma atualização de outro gateway. Observe que também há tempos de atualização para cada caminho. Quando o último caminho para um destino for removido, o destino é colocado em holddown, a menos que os holddowns estejam desativados (Consulte a seção Desativar holddowns para obter mais informações). A hora de validade do holddown indica a hora na qual o holddown expira. O fato de ser diferente de zero indica que o destino está em espera. Para economizar o tempo de cálculo, é bom manter uma métrica recomendada para cada destino. Isso é simplesmente o mínimo das métricas compostas para todos os caminhos até o destino.

Para cada caminho para um destino, há o endereço do próximo salto no caminho, a interface a ser usada, um vetor de métrica que caracteriza o caminho, incluindo retardo topológico, largura de banda, confiabilidade e ocupação do canal. Outras informações também estão associadas a cada caminho, incluindo contagem de saltos, MTU, fonte de informações, métrica composta remota e uma métrica composta calculada com base nesses números, de acordo com a equação 1. Há também uma hora de última atualização. A fonte de informações indica de onde provém a mais recente atualização para esse caminho. Na prática, isso é o mesmo que o endereço do próximo salto. O horário da última atualização é simplesmente a hora em que a atualização mais recente chegou nesse caminho. Ele é utilizado para caminhos com o intervalo de tempo esgotado.

Observe que a mensagem de atualização do IGRP tem três partes: interior, sistema (significando "este sistema autônomo" mas não interior) e exterior. A seção interna é para as rotas para sub-redes. Nem todas as informações de sub-rede estão incluídas. Somente as sub-redes de uma rede estão incluídas. É a rede associada ao endereço para o qual a atualização está sendo enviada. Normalmente, as atualizações são difundidas em cada interface, portanto, essa rede é a rede a partir da qual a difusão está sendo enviada. (Outros casos surgem por respostas a uma solicitação IGRP e IGRP de ponto a ponto.) Redes principais (por exemplo, que não são sub-redes) são colocadas na parte do sistema da mensagem de atualização, a menos que estejam especificamente sinalizadas como externas.

Uma rede será sinalizada como externa se foi aprendida de outro gateway e as informações chegaram na parte externa da mensagem de atualização. A implementação da Cisco também permite que o administrador do sistema declare redes específicas como exteriores. Rotas externas também são conhecidas como padrão candidato. São rotas que vão até os gateways ou passam pelos gateways considerados adequados como padrões, para serem usadas quando não houver rota explícita para um destino. Por exemplo, na Rutgers, configuramos o gateway que conecta a Rutgers à nossa rede regional, de modo que ela sinalize como externa a rota para o backbone NSFnet. A implementação do Cisco escolhe uma rota padrão selecionando a rota exterior com a menor métrica.

As seções seguintes esclarecem certas partes das Figuras 4 a 8.

Roteamento de Pacotes

A Figura 4 descreve o processamento total de pacotes de entrada. Seu uso tem por fim explicar a terminologia. Obviamente, esta não é uma descrição completa do que faz um gateway IP.

Esse processo utilize a lista de protocolos suportados e informações sobre as interfaces inseridas quando o gateway foi inicializado. Os detalhes do processamento do pacote dependem do protocolo usado pelo pacote. Isso é determinado na Etapa A. A Etapa A é a única parte da Figura 4 que é compartilhada por todos os protocolos. Quando o tipo de protocolo é conhecido, é usada a implementação da Figura 4 apropriada ao tipo de protocolo. Os detalhes do conteúdo do pacote são descritos pelas especificações do protocolo. As especificações de um protocolo incluem um procedimento para determinar o destino de um pacote, um procedimento para comparar o destino com os próprios endereços do gateway para determinar se o próprio gateway é o destino, um procedimento para determinar se um pacote é uma transmissão e um procedimento para determinar se o destino é parte de uma rede especificada. Esses procedimentos são usados nas etapas B e C da Figura 4. O teste na etapa D exige uma pesquisa dos destinos listados na tabela de roteamento. O teste será realizado se houver uma entrada na tabela de roteamento para o destino e esse destino estiver associado a ela em pelo menos um caminho utilizável. Observe que os dados de destino e do caminho usados nesta e na próxima etapa ficam separados para

cada tipo de serviço de suporte. Portanto, essa etapa começa com a determinação do tipo de serviço especificado pelo pacote e com a seleção do conjunto correspondente de estruturas de dados a ser usado para essa e para a próxima etapa.

Um caminho será útil para as etapas D e E se a métrica composta remota for menor que a métrica composta. Um caminho cuja métrica composta remota seja maior que a sua métrica composta, é um caminho cujo próximo salto é “mais distante” do destino, conforme medido pela métrica. Isso é conhecido como "caminho de upstream". Normalmente, era de se esperar que o uso de métricas evitasse que caminhos de upstream fossem escolhidos. É fácil perceber que um caminho de upstream nunca será o mais conveniente. No entanto, se uma grande variação é permitida, os caminhos não ideais podem ser usados. Alguns deles podem ser upstream.

O passo E computa o caminho a ser utilizado. Os caminhos cuja métrica composta remota não é menor que suas métricas compostas não são considerados. Se mais de um caminho é aceitável, eles são usados em uma forma ponderada de alternância circular (round robin). A freqüência com que um caminho é utilizado é inversamente proporcional à sua métrica composta.

Recebimento de Atualizações de Roteamento

A Figura 5 descreve o processamento de uma atualização de roteamento recebida de um gateway adjacente. Tais atualizações consistem em uma lista de entradas, cada uma com informações para um único destino. Pode haver mais de uma entrada para o mesmo destino em uma única atualização de roteamento, para acomodar vários tipos de serviços. Cada uma dessas entradas é processada individualmente, conforme descrito na Figura 5. Se uma entrada estiver na seção externa da atualização, o flag externo será definido para o destino se for adicionado como resultado do processo.

Todo o processo descrito na Figura 5 deve ser repetido uma vez para cada tipo de serviço suportado pelo gateway, usando o conjunto de informações de destino/caminho associadas a esse tipo de serviço. Isso é mostrado no loop mais externo na Figura 5. A atualização de roteamento inteira deve ser processada uma vez para cada tipo de serviço. Observe que a implementação atual de IGRP não suporta vários tipos de serviço, portanto o circuito mais externo não é de fato implementado.

No passo A, os testes básicos de aceitabilidade são feitos no caminho. Isso deve incluir testes de racionalidade para o destino. Os números de rede impossíveis ("marcianos") devem ser rejeitados. (Consulte RFC 1009 e RFC 1122 para obter mais informações.) As atualizações também são rejeitadas caso o destino ao qual elas se referem esteja suspenso, ou seja, o tempo de expiração da suspensão é diferente de zero e posterior à hora atual.

No Passo B, a tabela de roteamento é procurada para ver se esta entrada descreve um caminho que já é conhecido. Um caminho na tabela de roteamento é definido pelo destino ao qual ele está associado, pelo próximo salto listado como parte do caminho, pela interface de saída a ser usada para o caminho e pela fonte de informações (o endereço de origem da atualização, na prática, é o mesmo que o do próximo salto). A entrada do pacote de atualização descreve um caminho cujo destino está listado na entrada, cuja interface de saída é que originou a atualização e cujo salto seguinte e a fonte de informações são o endereço do gateway que enviou a atualização (a "origem" S).

Nas Etapas H e T, o processo de atualização descrito na Figura 7 é agendado. Na verdade, este processo será executado após a conclusão completa do processo descrito na figura 5. Ou seja, o processo de atualização descrito na Figura 7 ocorrerá apenas uma vez, mesmo se ele for

acionado várias vezes durante o processamento descrito na Figura 5. Além disso, são necessárias precauções para evitar a emissão de atualizações muito freqüentes quando a rede é alterada rapidamente.

A Etapa K é realizada se o destino descrito pela entrada atual no pacote de atualização já existe na tabela de roteamento. K compara a nova medição composta calculada a partir dos dados no pacote de atualização com a melhor medição composta do destino. Observe que a melhor métrica composta não é recalculada nesse momento, portanto, se o caminho considerado já estiver na tabela de roteamento, esse teste poderá comparar métricas novas e antigas para o mesmo caminho.

A Etapa L é realizada para os caminhos que estão piores do que a melhor métrica composta. Isso inclui novos caminhos que estão piores que os atuais e os caminhos atuais cuja métrica composta aumentou. A Etapa L testa se o novo caminho é aceitável. Observe que esse teste implementa o teste para saber se um novo caminho é bom o suficiente para ser mantido e o envenenamento de rota. Para ser aceitável, o valor do retardo não deve indicar destino inalcançável (para a implementação de IP atual, todos aqueles em um campo de 24 bits) e a métrica composta (calculada conforme especificado na figura 8) deve ser admissível. Para determinar se a métrica composta é aceitável, compare-a com a métrica composta de todos os demais caminhos até o destino. Deixe M ser o mínimo deles. O novo caminho será aceitável se for < V X M, ONDE V É A VARIANTE DEFINIDA QUANDO O GATEWAY FOI INCIALIZADO. SE V = 1 (O QUE SEMPRE É VERDADEIRO A PARTIR DO CISCO VERSÃO 8.2), ENTÃO QUALQUER UMA MÉTRICA PIOR QUE A EXISTENTE NÃO É ACEITÁVEL. HÁ UMA EXCEÇÃO PARA ISSO: SE O CAMINHO JÁ EXISTIR E FOR O ÚNICO CAMINHO PARA O DESTINO, O CAMINHO SERÁ MANTIDO SE A MÉTRICA NÃO PRECISAR SER AUMENTADA EM MAIS DE 10% (OU ONDE OS HOLDDOWNS ESTIVEREM DESATIVADOS, SE A CONTAGEM DE SALTOS NÃO TIVER AUMENTADO).

A etapa V está concluída quando as novas informações de um caminho indicam que a medição composta será reduzida. As métricas compostas de todos os caminhos até o destino D são comparada. Nesta comparação, a nova métrica composta para P é usada, em vez da que aparece na tabela de roteamento. O M mínimo da métrica composta é calculado. Em seguida, todos os caminhos para D são examinados novamente. Se a métrica composta para algum caminho for maior que M x V, o caminho é removido. V é a variação, inserida quando o gateway foi inicializado. (Desde o lançamento do Cisco 8.2, a variação está permanentemente definida como 1).

Processamento periódico

O processo descrito na Figura 6 é acionado uma vez por segundo. Ele examina vários temporizadores na tabela de roteamento para ver se algum deles expirou. Esses cronômetros são descritos acima.

Na Etapa U, o processo descrito na Figura 7 é ativado.

As Etapas R e S são necessárias porque as métricas compostas armazenadas na tabela de roteamento dependem da ocupação do canal que muda com o tempo, com base nas medições. A ocupação do canal é periodicamente recalculada, usando-se uma média móvel do tráfego medido através da interface. Se o valor recém-calculado diferir do valor existente, todas as métricas compostas envolvendo essa interface deverão ser ajustadas. Cada caminho mostrado na tabela de roteamento é examinado. Qualquer caminho cujo salto seguinte utilize a interface I possui sua métrica composta recalculada. Isso é feito de acordo com a Equação 1, utilizando como ocupação

do canal o máximo do valor armazenado na tabela de roteamenteo como parte da métrica de caminhos e a ocupação do canal recém calculada da interface.

Gerar mensagens de atualização

A Figura 7 descreve como o gateway gera mensagens de atualização para serem enviadas a outros gateways. Uma mensagem separada é gerada para cada interface de rede conectada ao gateway. Essa mensagem é enviada para todos os demais gateways que podem ser atingidos por meio da interface (Passo J). Em geral, isso é feito através do envio da mensagem como broadcast. Entretanto, se a tecnologia de rede ou o protocolo não permitir broadcasts, talvez seja necessário enviar a mensagem individualmente para cada gateway.

Em geral, a mensagem é criada ao adicionar uma entrada para cada destino na tabela de roteamento, na Etapa G. Observe que os dados de destino/caminho associados a cada tipo de serviço devem ser usados. Em último caso, uma nova entrada é adicionada à atualização de cada destino para cada tipo de serviço. Entretanto, antes de adicionar uma entrada na mensagem de atualização na Etapa G, as entradas já adicionadas são varridas. Se a nova entrada já está presente na mensagem de atualização, ela não é adicionada novamente. Uma nova entrada reproduz a atual quando os destinos e gateways de próximo salto são os mesmos.

Para simplificar, o pseudocódigo omite um item — as mensagens de atualização de IGRP têm três partes: interna, sistema e externa, o que significa que há realmente três loops de destinos. O primeiro inclui apenas as sub-redes da rede para a qual a atualização está sendo enviada. A segunda inclui todas as principais redes (por exemplo, as que não são sub-redes) não sinalizadas como externas. A terceira inclui todas as redes principais sinalizadas como externas.

O Passo E implementa o teste de horizonte dividido. No caso normal, esse teste falhará para rotas cujo melhor caminho passa pela mesma interface pela qual a interface está sendo enviada. No entanto, se a atualização estiver sendo enviada para um destino específico (por exemplo, em resposta a uma solicitação de IGRP, a partir de outro gateway ou como parte do IGRP ponto-a-ponto), o horizonte dividido falha somente se o melhor caminho, originalmente vindo daquele destino (sua origem de informação é a mesma do destino), e sua interface de saída forem iguais àquela de onde a solicitação veio.

Calcular informações sobre métrica

A figura 8 descreve como as informações de métricas são processadas a partir das mensagens atualizadas recebidas pelo gateway e como elas são geradas para que as mensagens atualizadas sejam enviadas pelo gateway. Observe que a entrada é baseada em um caminho em particular para o destino. Se houver mais de um caminho até o destino, um caminho com métrica composta mínima é escolhido. Se mais de um caminho tiver a métrica de composição mínima, uma regra arbitrária de quebra de vínculo será usada. (Para a maioria dos protocolos, isso se baseia no endereço do gateway de próximo salto).

Figura 4 — Processamento de pacotes recebidos

```
Data packet arrives using interface I

A   Determine protocol used by packet

    If protocol is not supported
      then discard packet
```

```
B   If destination address matches any of gateway's addresses
    or the broadcast address
      then process packet in protocol-specific way

C   If destination is on a directly-connected network
      then send packet direct to the destination, using
      the encapsulation appropriate to the protocol and link type

D   If there are no paths to the destination in the routing
    table, or all paths are upstream
      then send protocol-specific error message and discard the packet

E   Choose the next path to use. If there are more than
    one, alternate round-robin with frequency proportional
    to inverse of composite metric.

    Get next hop from path chosen in previous step.

    Send packet to next hop, using encapsulation appropriate
    to protocol and data link type.
```

Figura 5 — Processamento de atualizações de roteamento recebido

```
Routing update arrives from source S

    For each type of service supported by gateway
      Use routing data associated with this type of service

      For each destination D shown in update

A       If D is unacceptable or in holddown
          then ignore this entry and continue loop with next destination D

B       Compute metrics for path P to D via S (see Fig 8)

        If destination D is not already in the routing table
        then Begin

          Add path P to the routing table, setting last
          update times for P and D to current time.

H         Trigger an update

          Set composite metric for D and P to new composite
          metric computed in step B.

          End

        Else begin (dest. D is already in routing table)

K         Compare the new composite metric for P with best
          existing metric for D.

           New > old:

L            If D is shown as unreachable in the update,
             or holddowns are enabled and
              the new composite metric >
                (the existing metric for D) * V
                [use 1.1 instead of V if V = 1,
                as it is as of Cisco release 8.2]
O            or holddowns are disabled and
```

```
              P has a new hop count > old hop count
              then Begin

               Remove P from routing table if present

               If P was the last route to D
                 then Unless holddowns are disabled
                     Set holddown time for D to
                      current time + holddown time
T                   and Trigger an update

               End

              else Begin

               Compute new best composite metric for D

               Put the new metric information into the
               entry for P in the routing table

               Add path P to the routing table if it
               was not present.

               Set last update times for P and D to
               current time.

               End

            New <= OLD:

V             Set composite metric for D and P to new
              composite metric computed in step B.

              If any other paths to D are now outside the
              variance, remove them.

              Put the new metric information into the
              entry for P in the routing table

              Set last update times for P and D to
              current time.

           End

        End of for

      End of for
```

Figura 6 — Processamento periódico

```
Process is activated by regular clock, e.g. once per second

    For each path P in the routing table (except directly
    connected interfaces)

      If current time < P'S LAST UPDATE TIME + INVALID TIME
        THEN CONTINUE WITH THE NEXT PATH P

      Remove P from routing table

      If P was the last route to D
        then Set metric for D to inaccessible
          Unless holddowns are disabled,
```

```
            Start holddown timer for D and
          Trigger an update

        else Recompute the best metric for D

      End of for

    For each destination D in the routing table

      If D's metric is inaccessible
        then Begin

        Clear all paths to D

        If current time >= D's last update time + flush time
          then Remove entry for D

        End

      End of for

    For each network interface I attached to the gateway

R     Recompute channel occupancy and error rate

S     If channel occupancy or error rate has changed,
        then recompute metrics

      End of for

    At intervals of broadcast time

U     Trigger update
```

Figura 7 — Criação de atualização

```
Process is caused by "trigger update"

    For each network interface I attached to the gateway

      Create empty update message

      For each type of service S supported

        Use path/destination data for S

        For each destination D

E         If any paths to D have a next hop reached through I
            then continue with the next destination

          If any paths to D with minimal composite metric are
          already in the update message
            then continue with the next destination

G         Create an entry for D in the update message, using
          metric information from a path with minimal
          composite metric (see Fig. 8)

          End of for

        End of for
```

```
J     If there are any entries in the update message
       then send it out interface I

      End of for
```

Figura 8 — Detalhes da computação de métricas

Esta seção descreve o procedimento para computar métricas e contagens de saltos de uma atualização de roteamento de chegada. A entrada para esta função é a entrada para um destino específico em um pacote de atualização de roteamento. A saída é um vetor de métricas que pode ser usado para computar a métrica composta e uma contagem de saltos. Se este caminho for adicionado à tabela de roteamento, o vetor de métrica inteiro é inserido na tabela. Os parâmetros de interface usados nas seguintes definições são os definidos quando o gateway foi inicializado, para a interface em que a atualização de roteamento foi introduzida, com exceção de que a ocupação de canais e a confiabilidade têm como base uma média transitória de tráfego medido por meio da interface.

- Atraso = atraso do pacote + atraso topológico da interface
- Bandwidth = max (largura de banda do pacote, largura de banda do pacote)
- Reliability = min (confiabilidade a partir do pacote, confiabilidade da interface)
- Ocupação de canal = max (ocupação de canal do pacote, ocupação do canal de interface)(Máx. é usado para largura de banda, pois a métrica da largura de banda é armazenada na forma inversa. Conceitualmente, queremos a largura de banda mínima.) Observe que a ocupação original do canal do pacote deve ser salva, pois será necessário recomputar a ocupação efetiva do canal sempre que a ocupação do canal de interface for alterada.

Os elementos descritos a seguir não fazem parte do vetor métrico, mas também são mantidos na tabela de roteamento como características do caminho:

- Contagem de saltos = contagem de saltos do pacote.
- MTU = min (MTU a partir de pacote, MTU de interface).
- Métrica composta remota = calculada para equação 1 usando os valores de métrica do pacote. Isto é, os componentes de métrica são aqueles do pacote e não são atualizados como mostrado acima. Obviamente, isso deve ser calculado antes de terminar os ajustes mostrados acima.
- Métrica composta = calculada a partir da equação 1, utilizando valores métricos calculados como descrito nesta seção.

O resto desta seção descreve o procedimento de métricas de computação e contagem de nó das atualizações de roteamentos a serem enviadas.

Essa função determina as informações de métricas e a contagem de nós a serem colocadas em um pacote de atualização de saída. Baseia-se em um caminho específico para um destino, se houver quaisquer caminhos utilizáveis. Caso não haja caminhos ou os caminhos sejam todos de upstream, o destino é chamado de inacessível.

```
If destination is inaccessible, this is indicated by using a specific
        value in the delay field. This value is chosen to be larger
        than the largest valid delay. For the IP implementation this is
        all ones in a 24-bit field.

   If destination is directly reachable through one of the interfaces, use
        the delay, bandwidth, reliability, and channel occupancy of the
        interface. Set hop count to 0.
```

```
   Otherwise, use the vector of metrics associated with the path in the
        routing table. Add one to the hop count from the path in the
        routing table.
```

Detalhes da implementação de IP

Este resumo da seção descreve os formatos de pacote usados pelo Cisco IGRP. O IGRP é enviado usando datagramas IP com o protocolo IP 9 (IGP). O pacote é iniciado com um cabeçalho. Ele começa imediatamente após o cabeçalho IP.

```
unsigned version: 4;  /* protocol version number */
    unsigned opcode: 4;   /* opcode */
    uchar edition;        /* edition number */
    ushort asystem;       /* autonomous system number */
    ushort ninterior;     /* number of subnets in local net */
    ushort nsystem;       /* number of networks in AS */
    ushort nexterior;     /* number of networks outside AS */
    ushort checksum;      /* checksum of IGRP header and data */
```

Em mensagens de atualização, as informações de roteamento seguem imediatamente após o cabeçalho.

O número da versão é atualmente 1. Pacotes com outros números de versão são ignorados.

Opcode pode ser 1 = atualização ou 2 = requisição.

Isso indica o tipo de mensagem. O formato dos dois tipos de mensagem será mostrado abaixo.

Edição é um número de série que é incrementado sempre que há uma alteração na tabela de roteamento. (Isso é feito nas condições em que o pseudocódigo acima diz para acionar uma atualização de roteamento.) O número da edição permite que os gateways evitem as atualizações de processamento com informações que eles já viram. (Isso não está implementado atualmente.) Ou seja, o número da edição é gerado corretamente, mas ele será ignorado na entrada. Como é possível que pacotes sejam descartados, não está claro se o número da edição é suficiente para evitar o processo de duplicação. Seria necessário verificar se todos os pacotes associados à edição foram processados.

Asystem é o número de sistema autônomo. Na implementação da Cisco, um gateway pode participar de mais de um AS (Autonomous System, Sistema autônomo). Cada um desses sistemas executa seu próprio protocolo IGRP. Conceitualmente, existem tabelas de roteamento completamente separadas para cada sistema autônomo. As rotas que chegam de um sistema autônomo via IGRP são enviadas apenas em atualizações para esse AS. Este campo permite que o gateway selecione qual conjunto de tabelas de roteamento usar para processamento da mensagem. Se o gateway receber uma mensagem de IGRP para um AS não-configurado, a mensagem será ignorada. Na verdade, a implementação Cisco permite o "vazamento" de informações de um AS para outro. No entanto, eu considero isso uma ferramenta administrativa e não parte do protocolo.

Ninterior, nsystem e nexterior indicam o número de entradas em cada uma das três seções das mensagens de atualização. Essas seções foram descritas acima. Não há nenhuma outra demarcação entre as seções. As primeiras entradas ninterior são capturadas para serem internas, as próximas entradas nsystem como sistema e as últimas nexterior como externas.

A soma de verificação é de IP, calculada usando o mesmo algoritmo usado em uma soma de

verificação de UDP. A soma de verificação é calculada no cabeçalho do IGRP e em todas as informações de roteamento que o seguem. O campo de soma de verificação é definido como zero ao computá-la. A soma de verificação não inclui o cabeçalho IP, nem tem qualquer cabeçalho virtual, como o UDP e TCP.

Solicitações

Uma solicitação IGRP pede ao destinatário que envie sua tabela de roteamento. A mensagem de solicitação tem apenas um cabeçalho. Somente os campos de versão, opcode e um sistema são usados. Todos os outros campos são zero. Espera-se que o destinatário envie uma mensagem de atualização de IGRP normal ao requisitante.

Atualizações

Uma mensagem de atualização de IGRP contém um cabeçalho, seguido imediatamente por entradas de roteamento. São incluídas tantas entradas de roteamento quanto caberão em um datagrama de 1500 bytes (incluindo o cabeçalho IP). Com as declarações de estrutura atual, isso permite até 104 entradas. Se forem necessárias mais entradas, diversas mensagens de atualização são enviadas. Como mensagens de atualização são simplesmente processadas entrada por entrada, não há nenhuma vantagem em utilizar uma única mensagem fragmentada em vez de várias outras independentes.

Aqui está a estrutura de uma entrada de roteamento:

```
uchar number[3];       /* 3 significant octets of IP address */
    uchar delay[3];        /* delay, in tens of microseconds */
    uchar bandwidth[3];    /* bandwidth, in units of 1 Kbit/sec */
    uchar mtu[2];          /* MTU, in octets */
    uchar reliability;     /* percent packets successfully tx/rx */
    uchar load;            /* percent of channel occupied */
    uchar hopcount;        /* hop count */
```

Os campos definem que uchar[2] e uchar[3] são simplesmente inteiros binários de 16 e 24 bits, na ordem normal da rede de IP.

O número define o destino descrito. Ele é um endereço IP. Para economizar espaço, somente os 3 primeiros bytes do endereço IP são fornecidos, exceto na seção interior. Na seção interior, os últimos 3 bytes são fornecidos. Para rotas de sistema e externas, não são possíveis sub-redes, portanto o byte de ordem baixa é zero. Rotas internas são sempre sub-redes de uma rede conhecida, portanto, o primeiro byte desse número de rede é fornecido.

O retardo é em unidades de 10 microssegundos. Isto dá um intervalo de 10 microssegundos a 168 segundos, que parece suficiente. Um retardo geral indica que a rede não pode ser alcançada.

A largura de banda é a largura de banda inversa, em bits por segundos, em escala pelo fator 1.0e10. O intervalo é de uma linha de 1200 BPS a 10 Gbps. (Ou seja, se a largura de banda for N Kbps, o número usado é 10000000 / N).

O MTU está em bytes.

A confiabilidade é expressa como uma fração de 255. Isto é, 255 é 100%.

A carga é determinada como uma fração de 255.

A contagem de saltos é uma contagem simples.

A despeito das unidades um pouco estranhas usadas para largura de banda e retardo, alguns exemplos aparecem ordenados. Esses são os valores padrão usados para várias mídias comuns.

	Delay	Bandwidth
Satellite	200,000 (2 sec)	20 (500 Mbit)
Ethernet	100 (1 ms)	1,000
1.544 Mbit	2000 (20 ms)	6,476
64 Kbit	2000	156,250
56 Kbit	2000	178,571
10 Kbit	2000	1,000,000
1 Kbit	2000	10,000,000

Cálculos métricos

Aqui está uma descrição da forma como a métrica composta é realmente computada na versão do Cisco 8.0(3).

```
metric = [K1*bandwidth + (K2*bandwidth)/(256 - load) + K3*delay] *
                         [K5/(reliability + K4)]

    If K5 == 0, the reliability term is not included.

    The default version of IGRP has K1 == K3 == 1, K2 == K4 == K5 == 0
```

Informações Relacionadas

- Página de Suporte do IP Routing
- Página de suporte de IGRP
- Suporte Técnico - Cisco Systems