1 Protocolo IP (Internet Protocol)

1.1 – Arquitetura e Filosofia da Interligação em Redes

Conceitualmente, uma interligação em redes TCP/IP oferece três grupos de serviços, conforme a figura 1; sua distribuição na figura sugere que há dependências entre elas. No nível mais baixo, um serviço de transmissão sem conexão oferece um fundamento sobre o qual repousa tudo mais. No nível seguinte, um serviço de transporte confiável oferece uma plataforma de nível mais alto da qual dependem os aplicativos.

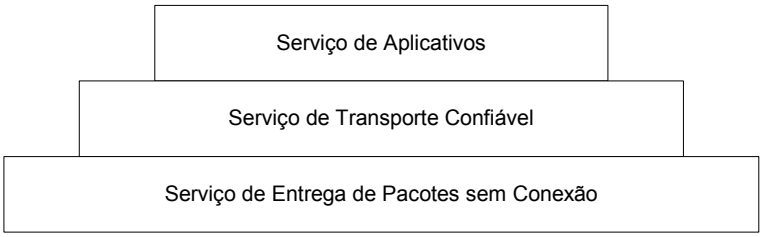

Figura 1

1.2 – O Conceito de Entrega Não-Confiável

Embora possamos associar softwares de protocolo a cada um dos serviços da figura 1, a razão de sua identificação como partes conceituais da interligação em redes é que elas claramente estabelecem as estruturas filosóficas do projeto. O ponto é:

> *O software da interligação em redes é projetado com base em três serviços de rede conceituais organizados hierarquicamente; grande parte de seu sucesso ocorreu porque essa arquitetura é surpreendentemente eficiente e adaptável.*

Uma das vantagens mais significativas dessa divisão conceitual é que ela tona possível substituir um serviço sem prejudicar os demais. Assim, a pesquisa e o desenvolvimento podem progredir simultaneamente ao longo dos três serviços.

1.3 – Sistema de Transmissão sem Conexão

O serviço mais importante da interligação em redes consiste em um sistema de entrega de pacotes. Tecnicamente, o serviço é definido como um sistema de transmissão sem conexão, *best-effort* e não confiável; é análogo ao serviço oferecido por hardwares de redes que operam em um paradigma de transmissão *best-effort*. O serviço é conhecido como *não-confiável* porque a entrega não é garantida. O pacote pode ser perdido, reproduzido, atrasar-se ou ser entregue com problemas, mas o serviço não detectará tais condições, nem informará isso ao transmissor nem ao receptor. Ele é denominado *sem conexão* porque cada pacote é independente dos outros. Uma seqüência de pacotes enviados de um computador a outro pode trafegar por caminhos diferentes, ou alguns podem ser perdidos enquanto outros são entregues. Finalmente, o serviço utiliza uma *transmissão best-effort* porque o software de

interligação em redes faz uma série de tentativas para entregar os pacotes. Isso significa que a interligação em redes não rejeita pacotes por simples capricho; a não-confiabilidade surge quando os recursos esgotam-se ou as redes básicas falham.

1.4 – Finalidade do Protocolo de Interligação em Redes

O protocolo que define o mecanismo de transmissão sem conexão e não-confiável é conhecido como In*ternet Protocol*, e é normalmente citado por suas iniciais *IP*. O IP oferece três definições importantes. Primeira, o protocolo IP define a unidade básica de transferência de dados utilizada através de uma interligação em redes TCP/IP. Assim, ela especifica o formato exato de todos os dados à medida que ela passa pela interligação em redes TCP/IP. Segunda, o software IP desempenha a função de *roteamento*, escolhendo um caminho por onde os dados serão enviados. Terceira, além da especificação formal e precisa de formatos de dados e de roteamento, o IP inclui um conjunto de regras que concentram a idéia da entrega não-confiável de pacotes. As regras definem como os hosts e os roteadores devem processar os pacotes, como e quando as mensagens de erro devem ser geradas e as condições segundo as quais os pacotes podem ser descartados. O IP é uma parte tão fundamental do projeto de uma interligação em redes TCP/IP às vezes é denominada uma *tecnologia baseada em IP*.

1.5 – O Datagrama de Interligação em Redes

A analogia entre uma rede física e uma interligação em redes TCP/IP é grande. Numa rede física, a unidade de transferência é um quadro que contém um cabeçalho e dados, onde o cabeçalho fornece informações como endereço de origem e de destino (físico). A interligação em redes denomina sua unidade básica de transferência de um *datagrama de interligação em redes*, às vezes citado como um *datagrama IP*, ou simplesmente um *datagrama*. Como um quadro típico de rede física, um datagrama é dividido em cabeçalho e áreas de dados. Também como um quadro, o cabeçalho de um datagrama contém os endereços de origem e destino e um tipo de campo que identifica o conteúdo do datagrama. Naturalmente, a diferença é que o cabeçalho do datagrama contém os endereços IP, enquanto o quadro contém os endereços físicos. A figura 2 mostra o formato geral de uma datagrama.

Figura 2

1.6 – Formato do Datagrama

Agora que já descrevemos o formato geral de um datagrama IP, podemos examinar o conteúdo com mais detalhes. A figura 3 mostra a organização dos campos em um datagrama.

Já que o processamento de datagramas se dá em softwares, o conteúdo e o formato não são restringidos por quaisquer hardwares. O primeiro campo de quatro bits de um datagrama (VERS), por exemplo, contém a versão do protocolo IP utilizada para criar o datagrama. Ele é utilizado para verificar se o transmissor, o receptor e quaisquer roteadores existentes entre ele concordam quanto ao formato

do datagrama. Todo software IP precisa verificar o campo de versão antes de processar um datagrama, para assegurar-se de que ele se adapta ao formato que o software espera. Se os padrões mudarem, as máquinas rejeitarão datagramas com versões de protocolos diferentes dos seus, impedindo que eles deturpem o conteúdo da datagrama com um formato desatualizado. A versão atual do protocolo IP é a quatro.

O campo de comprimento do cabeçalho (HLEN), também de quatro bits, fornece o comprimento do cabeçalho do datagrama medido em palavras de 32 bits. Como veremos, todos os campos do cabeçalho contêm um comprimento fixo, exceto para OPÇÕES IP e os campos correnpondentes PADDING. O cabeçalho mais comum, que não contém qualquer opção e nenhum preenchimento, mede 20 octetos e o campo de comprimento de cabeçalho é cinco.

O campo COMPRIMENTO TOTAL fornece o comprimento do datagrama IP medido em octetos, incluindo octetos no cabeçalho e nos dados. O tamanho da área de dados pode ser calculando subtraindo-se de COMPRIMENTO TOTAL o comprimento do cabeçalho (HLEN). Já que o campo COMPRIMENTO TOTAL possui 16 bits de comprimento, o maior tamanho possível para um datagrama IP é 2^{16} ou 65-535 octetos. Na maioria dos aplicativos, essa não é uma limitação rígida. No futuro pode tornar-se mais importante, se as redes de velocidade mais alta puderem transportar pacotes de dados maiores que 65.535 octetos.

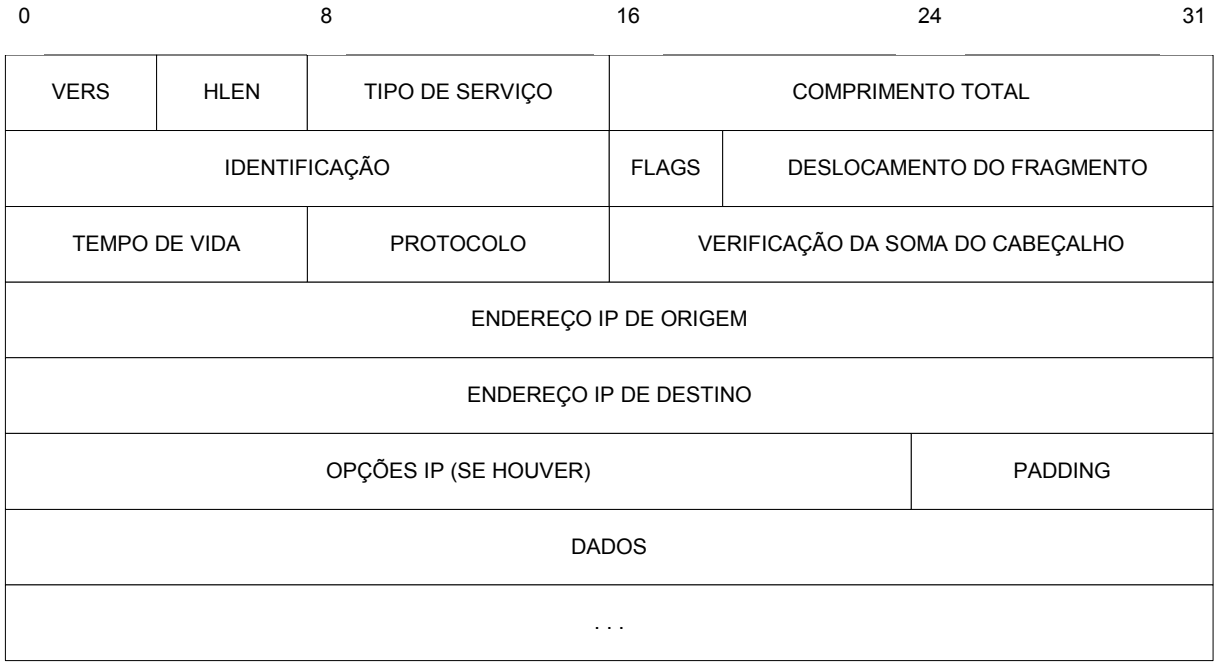

Figura 3

1.7 – Os Campos do Tipo de Serviço e Precedência

Denominado informalmente *Type of Service* (*TOS*), o campo TIPO DE SERVIÇO, de oito bits, especifica como o datagrama deve ser tratado e é fracionado em cinco subcampos, como mostra a figura 4.

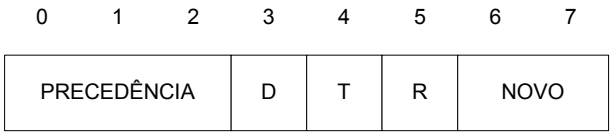

Figura 4

Três bits PRECEDÊNCIA especificam a precedência do datagrama com valores variando de zero (precedência normal) até sete (controle de rede), permitindo que os transmissores indiquem a importância de cada datagrama. Embora a maioria dos softwares que rodam e hosts ignorem o tipo de serviço, trata-se de um conceito importante, porque fornece um mecanismo que pode permitir que informações de controle tenham precedência sobre dados. Se, por exemplo, todos os hosts e roteadores reconhecem a precedência, é possível implementar algoritmos de controle de congestionamento que não sejam influenciados pelo congestionamento que estão tentando controlar.

Os bits D, T e R especificam o tipo de transporte que o datagrama deseja. Quando ajustado, o bit D solicita um intervalo baixo, o bit T solicita um throughput alto e o bit R solicita alta confiabilidade. É claro que não deve ser possível que a interligação em redes garanta o tipo de transporte solicitado (ou seja, pode acontecer que nenhum caminho para o destino tenha a propriedade solicitada). Assim, consideramos, a solicitação de transporte como uma sugestão para os algoritmos de roteamento, e não uma exigência. Se um roteador realmente conhece mais de uma rota possível para determinado destino, ele pode utilizar o tipo de campo de transporte para selecionar aquelas cujas características mais se aproximem das desejadas. Suponha, por exemplo, que um roteador possa selecionar entre uma linha alugada, de baixa capacidade, e uma conexão de alta, de satélites de banda larga (mas de intervalo alto). O conjunto de bits D poderia solicitar aos datagramas que carregam toque no teclado de um usuário para um computador remoto que esses sejam entregues o mais rápido possível, enquanto um conjunto de bits T poderia solicitar aos datagramas correspondentes uma transferência de arquivos que trafeguem nos links de alta capacidade de satélites.

Também é importante entender que os algoritmos de roteamento precisam escolher entre tecnologias de redes físicas básicas as quais possuem, cada uma, características de intervalo, throughput e confiabilidade. De uma maneira geral, algumas tecnologias representam um compromisso entre duas características (por exemplo, um maior throughput em detrimento de maiores retardos). Assim, a idéia é apresentar uma sugestão ao algoritimo de roteamento sobre o que é mais importante, e raramente faz sentido especificar os três tipos de serviço. Para resumir:

> *Consideramos o tipo de especificação de transporte uma sugestão ao algoritmo de roteamento para ajudá-lo a escolher entre os vários caminhos para um destino, com base em seu conhecimento das tecnologias de hardware disponíveis nesses caminhos. Uma interligação em redes não garante o tipo de transporte solicitado.*

1.8 – Encapsulamento de Datagramas

Antes que possamos compreender os próximos campos de um datagrama, é importante considerar como eles relacionam-se com os quadros de redes físicas. Começaremos com uma pergunta: “que tamanho um datagrama pode ter?” Ao contrário de quadros de redes físicas que precisam ser reconhecidos pelo hardware, os datagramas são tratados por software. Eles podem ter qualquer tamanho que os projetistas de protocolos escolherem. Já vimos que o atual formato de datagrama aloca somente 16 bits para o campo de comprimento total, limitando o datagrama a, no máximo, 65.535

octetos. Entretanto, esse limite poderia ser mudado em versões posteriores do protocolo. Os limites mais importantes para o tamanho de datagramas surgem na prática. Sabemos que, à medida que os datagramas movem-se de uma máquina para outra, eles precisam sempre ser transportados por uma rede física básica. Para tornar o transporte da interligação em redes eficiente, gostaríamos de assegurar que cada datagrama viaje em um quadro físico distinto. Isso significa que desejamos que nossa abstração de um pacote de rede física mapeie diretamente para dentro de um pacote real, se possível.

A idéia de transportar um datagrama em um quadro de rede é denominada *encapsulamento*. Para a rede básica, um datagrama é como qualquer outra mensagem enviada de uma máquina a outra. O hardware não reconhece o formato do datagrama, nem entende o endereço de destino IP. Assim, conforme mostra a figura 5, quando uma máquina envia um datagrama IP a outra, todo o datagrama é transportado na porção de dados do quadro de rede.

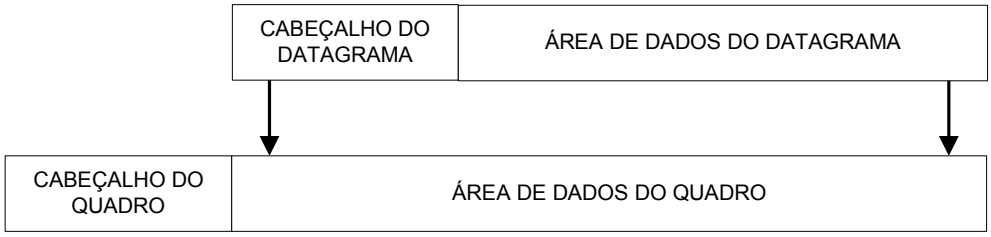

Figura 5

1.9 – Tamanho do Datagrama, MTU da Rede e Fragmentação

Na situação ideal, todo o datagrama IP encaixa-se em um quadro físico, tornando a transmissão na rede física eficiente. Para obter essa eficiência, os projetistas de IP devem ter selecionado um tamanho máximo de datagrama, de modo que ele sempre se encaixe em um quadro. Mas que tamanho de quadro deve ser escolhido? Acima de tudo, um datagrama pode trafegar em muitos tipos de redes físicas, à medida que move-se na interligação em redes para o seu destino final.

Para compreender o problema, precisamos entender um fato sobre o hardware de rede: cada tecnologia de comutação de pacotes coloca um limite superior, fixo, no total de dados que podem ser transferido em um quadro físico. A Ethernet, por exemplo, limita as transferências a 1.500 octetos de dados, enquanto a FDDI permite aproximadamente 4.470 octetos de dados por quadro. Referimo-nos a esses limites como MTU (m*aximum transfer unit*). O tamanho da MTU pode ser bem pequeno: algumas tecnologias de hardware limitam a transferência para 128 octetos ou menos. Limitar os datagramas para encaixar a menor MTU possível na interligação em redes torna a transferência ineficaz quando aqueles datagramas trafegam em uma rede que pode transportar quadros de tamanho maiores. Enquanto, permitir que os datagramas sejam maiores que a MTU mínima da rede em uma interconexão significa que um datagrama nem sempre irá encaixar-se no quadro único de uma rede.

A escolha deve ser óbvia: o objetivo do projeto de interligação em redes é concentrar as tecnologias de rede básicas e facilitar a comunicação para o usuário. Assim, em vez de projetar datagramas que sigam as restrições de redes físicas, o software TCP/IP escolhe um tamanho inicial de datagrama conveniente e descobre uma forma de dividir os datagramas extensos em frações menores, quando o datagrama precisa atravessar uma rede que tenha uma MTU pequena. As pequenas frações em que um datagrama é dividido são denominada *fragmentos*, e o processo de divisão de um datagrama é conhecido como *fragmentação*.

Conforme a figura 6, a fragmentação normalmente ocorre em um roteador situado em algum ponto ao longo do caminho entre a origem do datagrama e seu destino final. O roteador recebe um datagrama de uma rede com uma MTU grande, e precisa enviá-lo em uma rede para a qual a MTU seja menor do que o tamanho do datagrama.

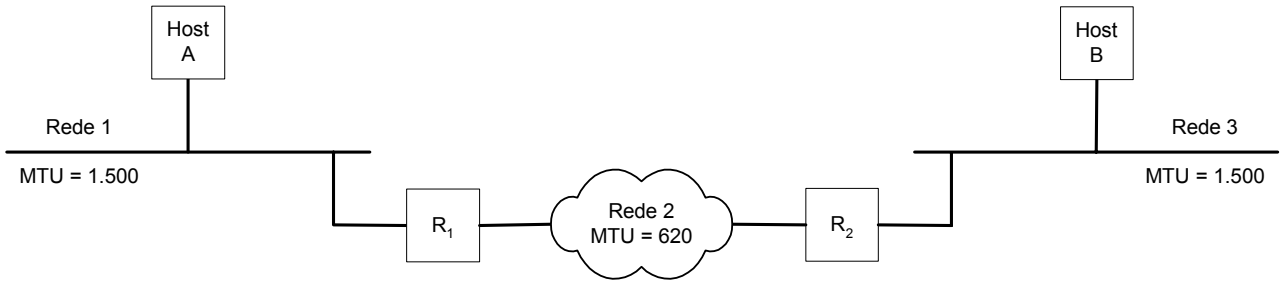

Figura 6

Na figura, ambos os hosts conectam-se diretamente às Ethernets que possuam uma MTU de 1.500 octetos. Assim, os dois hosts podem gerar e enviar datagramas de até 1.500 octetos de comprimento. O caminho entre eles, entretanto, inclui uma rede com uma MTU de 620. Se o host A envia ao host B um datagrama maior que 620 octetos, o roteador R_1 fragmentará o datagrama. Da mesmo forma, se B envia um datagrama grande a A, o roteador R_2 fragmentará o datagrama.

O tamanho do fragmento é escolhido de tal forma que cada fragmento possa ser transportado na rede básica em um quadro único. Além disso, já que o IP representa o deslocamento dos dados em múltiplos de oito octetos, o tamanho do fragmento precisa ser um múltiplo de oito. É claro que escolher o múltiplo de oito octetos mais próximo do MTU da rede nem sempre divide o datagrama em frações de igual tamanho; a última fração é normalmente menor que as outras. Os fragmentos devem ser *remontados* para produzir uma cópia completa do datagrama original, antes que ele possa ser processado no destino.

O protocolo IP não limita datagramas a um tamanho pequeno, nem garante que datagramas grandes serão entregues sem fragmentação. A origem pode escolher qualquer tamanho de datagrama que julgar apropriado; a fragmentação e remontagem ocorrem automaticamente, sem qualquer ação específica por parte da origem. A especificação do IP indica que os roteadores precisam aceitar datagramas até o máximo de MTUs de rede às quais se conectam. Além disso, um roteador precisa sempre tratar os datagramas de até 576 octetos. (Hosts também devem aceitar e remontar, se necessário, os datagramas de, no mínimo, 576 octetos.)

Fragmentar um datagrama significa dividi-lo em várias frações. Deve surpreendê-lo o fato de que cada fração tem o mesmo formato que o datagrama original. A figura 7 ilustra o resultado da fragmentação.

Cada fragmento contém um cabeçalho de datagrama que duplica a maior parte do cabeçalho do datagrama original (exceto para um bit no campo FLAGS que mostra que é um fragmento), seguido por tantos dados quantos puderem ser transportados no fragmento, enquanto mantém o comprimento total menor que a MTU da rede na qual precisa trafegar.

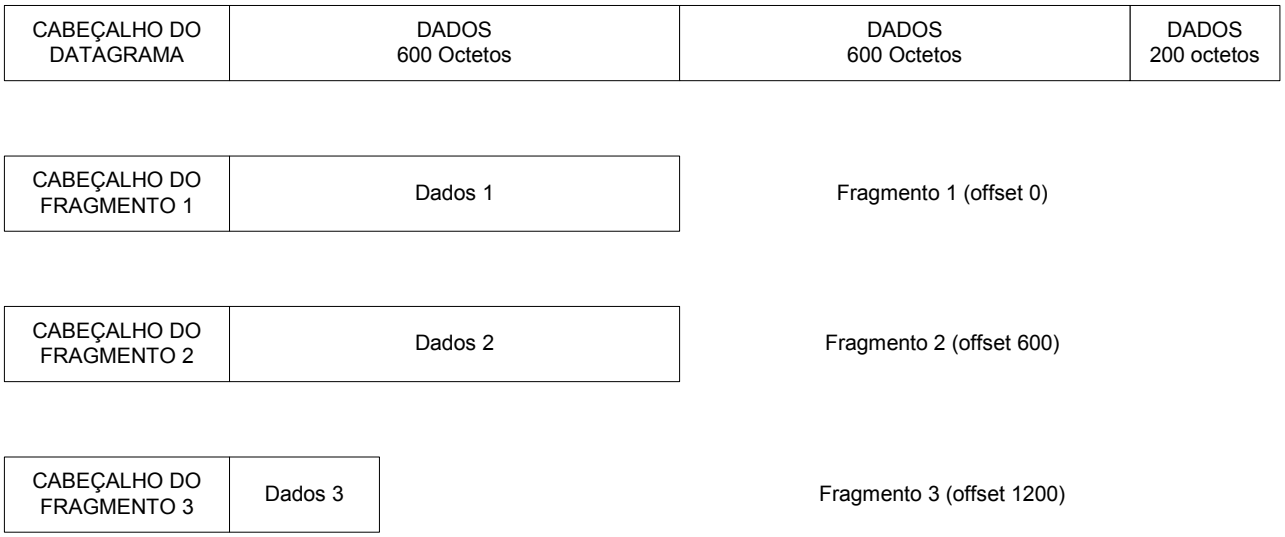

Figura 7

1.10 – Remontagem de Fragmentos

O datagrama deve ser remontado após passar em uma rede, ou os fragmentos devem ser transportados para o host final antes da remontagem? Em uma interligação de redes TCP/IP, quando um datagrama tiver sido fragmentado, os fragmentos trafegam como datagramas isolados ao longo do percurso até o último destino onde precisam ser remontados. Há duas desvantagens em preservar os fragmentos ao longo do percurso até o final. Primeira, porque datagramas não são remontados imediatamente após passarem por uma rede com uma MTU pequena, e os fragmentos pequenos precisam ser transportados do ponto de fragmentação até o destino final. A remontagem de datagramas no destino final pode levar à ineficiência: mesmo se algumas redes físicas encontradas após o pronto de fragmentação possuírem grande capacidade de MTU, apenas pequenos fragmentos atravessam-na. Segunda, se quaisquer fragmentos forem perdidos, o datagrama não pode ser remontado. A máquina receptora inicia um *temporizador de remontagem* quando recebe um fragmento inicial. Se o temporizador terminar antes que todos os fragmentos cheguem, a máquina receptora descarta os fragmentos remanescentes sem processar o datagrama. Assim, a probabilidade de perda de datagrama cresce quando a fragmentação ocorre, porque a perda de um fragmento único resulta na perda do datagrama inteiro.

Apesar das pequenas desvantagens, a execução de remontagem no destino final funciona bem. Permite que cada fragmento seja roteado independentemente, e não exige que roteadores intermediários armazenem ou remontem fragmentos.

1.11 – Controle de Fragmentação

Três campos no cabeçalho do datagrama, IDENTIFICAÇÃO, FLAGS e OFF-SET DE FRAGMENTO, controlam a fragmentação e a remontagem de datagramas. O campo IDENTIFICAÇÃO contém um número inteiro único que identifica o datagrama. Lembre-se de que, quando um roteador fragmenta um datagrama, ele copia a maioria dos campos no cabeçalho do datagrama para cada fragmento. O campo IDENTIFICAÇÃO precisa ser copiado. Sua finalidade é permitir que o destino saiba quais datagramas

estão chegando e a que datagramas pertencem. À medida que chega um fragmento, o destino utiliza o campo IDENTIFICAÇÃO juntamente com o endereço de origem do datagrama para que esse seja identificado. Os computadores que estão enviando os datagramas IP devem gera um valor único para o campo IDENTIFICAÇÃO, para cada datagrama. Uma técnica utilizada pelo software IP mantém uma contagem global em memória, incrementa-a a cada vez que um novo datagrama é criado e atribui o resultado como o campo IDENTIFICAÇÃO do datagrama.

Lembre-se de que cada fragmento possui exatamente o mesmo formato que um datagrama completo. Para um fragmento, o campo OFFSET DE FRAGMENTO especifica o deslocamento, no datagrama original, dos dados que estão sendo transportados no fragmento, medidos em unidades de oito octetos, iniciando em deslocamento zero. Para remontar o datagrama, o destino precisa obter todos os fragmentos que iniciam com o framento que possui deslocamento zero até o fragmento de maior deslocamento. Os fragmentos não chegam necessariamente em ordem, e não há comunicação entre o roteador que fragmento o datagrama e o destino que está tentando remontá-lo.

Os dois bits de baixa ordem, do campo FLAGS de três bits, controlam a fragmentação. Normalmente, o software aplicativo que utiliza TCP/IP não dá atenção à fragmentação, porque essa e a remontagem são procedimentos automáticos que ocorrem em um baixo nível do sistema operacional, invisível para usuários finais. Entretanto, para testar o software de interligação em redes ou depurar problemas operacionais, deve ser importante testar os tamanhos dos datagramas para os quais a fragmentação ocorre. O primeiro bit de controle auxilia nesse teste, especificando se o datagrama pode ser fragmentado. Ele é conhecido como bit *não-fragmentar* porque o seu ajuste em um especifica que o datagrama não deve ser fragmentado. Um aplicativo pode optar por não permitir uma fragmentação quando somente o datagrama inteiro é útil. Considere, por exemplo, um procedimento de inicialização de um computador, no qual uma máquina começa a executar um pequeno programa na ROM que utiliza a interligação em redes para solicitar uma inicialização inicial, e uma outra máquina retorna uma imagem de memória. Se o software tiversido projetado de tal modo que a imagem de inicialização só tenha utilidade se obtida de uma única vez, o datagrama deve ter um conjunto de bits *não-fragmentar*. Toda vez que um roteador precisa fragmentar um datagrama que possui o conjunto de bits *não-fragmentar*, o roteador descarta o datagrama e retorna à origem uma mensagem de erro.

O bit de mais baixa ordem, no campo FLAGS, especifica se o fragmento contém a parte do meio ou do final dos dados do datagrama. Ele é conhecido como bit de *mais fragmentos*. Para verificar por que esse bit é necessário, considere o software IP no destino final tentando remontar um datagrama. Ele receberá fragmentos (possivelmente fora de ordem) e precisa saber quando recebeu todos os fragmentos para um datagrama. Quando um fragmento chega, o campo COMPRIMENTO TOTAL, do cabeçalho, aplica-se ao tamanho do fragmento e não ao tamanho do datagrama original; assim, o destino não pode utilizar o campo COMPRIMENTO TOTAL para inferir se reuniu todos os fragmentos. O bit de *mais fragmentos* resolve o problema facilmente: quando o destino recebe um fragmento com o bit *mais fragmentos* desativado, ele sabe que o esse fragmento transporta a parte final dos dados do datagrama original. Partindo dos campos de OFFSET DE FRAGMENTO e COMPRIMENTO TOTAL, ele pode calcular o comprimento do datagrama original. Examinando o OFFSET DE FRAGMENTO e o COMPRIMENTO TOTAL de todos os fragmentos que chegaram, um receptor pode dizer se os fragmentos sob controle contêm todos os dados necessários para remontar todo o datagrama original.

1.12 – TTL (Time To Live ou Tempo de Vida)

O campo TEMPO DE VIDA especifica quanto tempo, em segundos, o datagrama pode permanecer no sistema de interligação em redes. A idéia é simples e importante: toda vez que uma máquina injeta um datagrama na interligação em redes, ela estabelece um tempo máximo de vida para o datagrama. Os

roteadores e os hosts que processam datagramas precisam decrementar o campo TEMPO DE VIDA à medida que o tempo passa e remover o datagrama da interligação em redes quando seu tempo expira.

Estimar o tempo exato é difícil porque os roteadores geralmente não sabem o tempo de trânsito para redes físicas. Poucas regras simplificam o processamento e facilitam o tratamento dos datagramas em o uso de relógios sincronizados. Primeiramente, cada roteador colocado ao longo do trajeto, da origem ao destino, precisa decrementar em um campo TEMPO DE VIDA quando ele processa o cabeçalho do datagrama. Além disso, para tratar as ocorrências de roteadores sobrecarregados implicam retardos longos, cada roteador registra o tempo local quando o datagrama chega e decrementa o TEMPO DE VIDA no número de segundos que o datagrama permaneceu dentro do roteador esperando serviço.

Sempre que um campo TEMPO DE VIDA alcança zero, o roteador descarta o datagrama e envia uma mensagem de erro de volta à origem. A idéia de manter um temporizador para datagramas é interessante, porque assegura que os datagramas não podem trafegar indefinidamente na interligação em redes, ainda que as tabelas de roteamento fiquem destruídas e os roteadores direcionem datagramas em círculo.

1.13 – Outros Campos do Cabeçalho de Datagramas

O campo PROTOCOLO é análogo ao campo de tipo em um quadro de rede. O valor no campo PROTOCOLO especifica qual protocolo de alto nível foi utilizado para criar a mensagem que está sendo transportada na área de DADOS do datagrama. Na verdade, o valor PROTOCOLO especifica o formato da área DADOS. O mapeamento entre um protocolo de alto nível e o valor de número inteiro no campo PROTOCOLO precisa ser administrado por uma autoridade central para garantir um consenso em toda a Internet.

O campo VERIFICAÇÃO DA SOMA DO CABEÇALHO assegura a integridade dos valores de cabeçalho. A verificação IP é formada com o tratamento do cabeçalho como uma seqüência de números inteiros de 16 bits (na ordem de bytes da rede), reunindo-os com uma aritmética complemento de um, e a seguir considerando o complemento de um como o resultado. Para a finalidade de calcular a soma de verificação, considera-se que o campo VERIFICAÇÃO DA SOMA DO CABEÇALHO contenha 0.

É importante observar que a soma de verificação somente se aplica a valores do cabeçalho IP, e não aos dados. Há vantagens e desvantagens em separar a soma de verificação para cabeçalhos e dados. Como o cabeçalho normalmente ocupa menos octetos que os dados, ter uma soma de verificação separada reduz o tempo de processamento nos roteadores que somente precisam calcular somas de verificação de cabeçalhos. A separação também permite que protocolos de mais alto nível escolham seu próprio esquema de soma de verificação para os dados. A principal desvantagem é que os protocolos de mais alto nível são forçados a acrescentar sua própria soma de verificação, ou arriscar que dados destruídos prossigam sem que sejam detectados.

Os campos ENDEREÇO IP DE ORIGEM e ENDEREÇO IP DE DESTINO contêm endereços IP de 32 bits do transmissor do datagrama e do receptor desejado. Embora o datagrama possa ser roteado através de muitos roteadores intermediários, os campos da origem e destino nunca mudam; eles especificam os endereços IP da origem e do último destino.

O campo denominado DADOS na figura 6 mostra o início da área de dados do datagrama. Seu comprimento depende, logicamente, do que está sendo enviado no datagrama. O campo OPÇÕES IP, abordado abaixo, é de comprimento variável. O campo denominado PADDING depende das opções selecionadas. Ele representa bits contendo zero e que podem ser necessários para garantir que o

cabeçalho do datagrama se estenda até o múltiplo exato de 32 bits (lembre-se de que o campo de comprimento do cabeçalho é especificado em unidades de palavra de 32 bits).

1.14 – Opções nos Datagramas de Interligação em Redes

O campo OPÇÕES IP que se segue ao endereço de destino não é necessário em todo datagrama, e as opções são incluídas principalmente para testes ou depuração da rede. Contundo, o processamento de opções é parte integrante do protocolo IP; assim, todas as implementações de padrões precisam incluí-lo.

O comprimento do campo OPÇÕES IP varia de acordo com as opções selecionadas. Algumas delas têm um octeto de comprimento e consistem em um único *código de opção* de octeto. Outras opções têm comprimento variável. Quando as opções são apresentadas em um datagrama, elas aparecem bem próximas, sem quaisquer separadores especiais. Cada opção consiste em um código de opção de octeto único que pode ser seguido por um comprimento de octeto único e um conjunto de octetos de dados para aquela opção. O octeto de código de opção é dividido em três campos, conforme a figura 8.

Os campos consistem em um flag CÓPIA de um bit, uma CLASSE DE OPÇÕES de dois bits e um NÚMERO DE OPÇÕES de cinco bits. O flag CÓPIA controla o modo como os roteadores tratam as opções durante a fragmentação. Quando o bit CÓPIA é ajustado em um, ele especifica que a opção deve ser copiada em todos os fragmentos. Quando ajustado em zero, o bit CÓPIA significa que a opção somente deve ser copiada no primeiro fragmento, e não em todos eles.

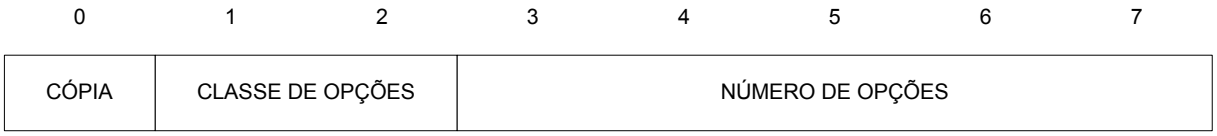

Figura 8

Os bits CLASSE DE OPÇÕES e NÚMERO DE OPÇÕES especificam a classe geral da opção e fornecem uma opção específica nessa classe. A tabela da figura 9 mostra como as classes são atribuídas.

Classes de Opções	Significado
0	Controle de rede ou de datagrama
1	Reservado para utilização futura
2	Depuração e avaliação
3	Reservado para utilização futura

Figura 9

A tabela da figura 10 relaciona as opções possíveis que podem acompanhar um datagrama IP e fornece seus valores CLASSE DE OPÇÕES e NÚMERO DE OPÇÕES. Conforme mostra a lista, a maioria das opções é utilizada para fins de controle.

Classe de Opções	Número de Opções	Comprimento	Descrição
0	0	-	Fim da lista de opções. Será utilizado se opções terminarem no fim do cabeçalho (ver também campo padding do cabeçalho).
0	1	-	Nenhuma operação (utilizado para alinhar actetos em um alista de opções).
0	2	11	Restrições de segurança e tratamento (para aplicações militares).
0	3	Var	Roteamento de origem separado. Utilizado para rotear um datagrama ao longo de um caminho específico.
0	7	Var	Rota de registro. Utilizado para traçar uma rota.
0	8	4	Identificador de fluxo. Utilizado para transportar um identificador de fluxo SATNET (Obsoleto).
0	9	Var	Roteamento de origem restrito. Utilizado para rotear um datagrama ao longo de um caminho especificado.
2	4	Var	Indicação de hora da inter-rede. Utilizado para registrar a indicação de hora ao longo da hora.

Figura 10

1.15 – Opção de Armazenamento de Rota

As opções de indicação de hora (timestamp) e de roteamento são as mais interessantes porque oferecem uma forma para monitorar ou controlar como os roteadores de interligação em redes direcionam datagramas. A opção rota de registro permite que a origem crie uma lista vazia de endereços IP e faz com que o endereço IP de cada roteador que processe o datagrama seja acrescentado à lista. A figura 11 mostra o formato da opção de rota de registro.

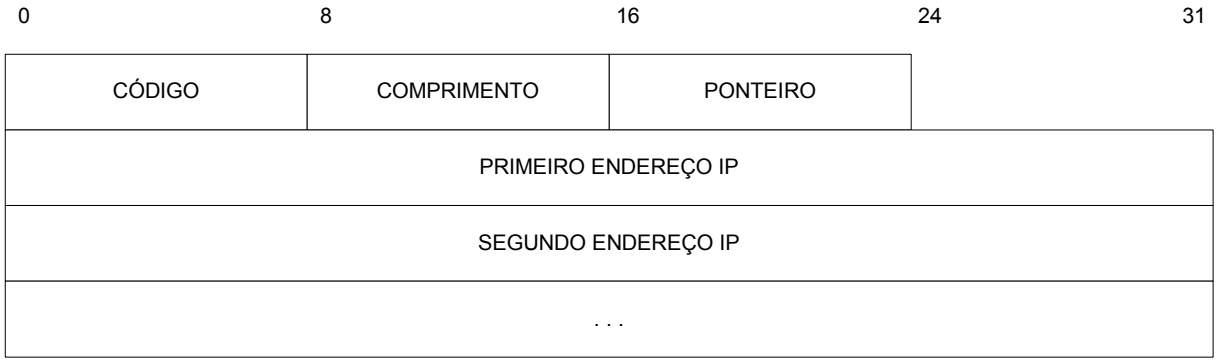

Figura 11

Conforme foi anteriormente descrito, o campo CÓDIGO contém a classe de opções e o número de opções (zero e sete para a rota de registro). O campo COMPRIMENTO especifica o comprimento total da opção conforme aparece no datagrama IP, inclusive os três primeiros octetos. Os campos que iniciam com o rótulo PRIMEIRO ENDEREÇO IP compreendem uma área reservada para a gravação de endereços da interligação em redes. O campo PONTEIRO especifica o deslocamento dentro da opção do próximo slot disponível.

Toda vez que uma máquina processa um datagrama. No qual é setada a opção de armazenamento da rota, ela acrescenta seu endereço à lista (espaço suficiente deve ser alocado pela origem para manter todas as entradas necessárias). Para acrescentar-se à lista, uma máquina primeiro compara os campos de ponteiro e comprimento. Se o ponteiro for maior que o comprimento, a lista estará completa. De modo que a máquina encaminhe o datagrama sem inserir sua entrada. Se a lista não estiver completa, a máquina insere seu endereço IP de quatro octetos na posição especificada pelo PONTEIRO e incrementa o ponteiro quatro posições.

Quando o datagrama chega, a máquina de destino pode extrair e processar a lista de endereços IP. Normalmente, um computador que recebe um datagrama ignora a rota armazenada. Utilizar a opção de rota de registro exige que duas máquinas concordem em cooperar; um computador não irá automaticamente receber rotas armazenadas em datagramas que saem. A origem precisa concordar em ativar a opção de rota de registro e do destino precisa concordar em processar a lista resultante.

1.16 – Opções de Rota de Origem

Uma outra idéia que outros projetista de rede acham interessante é a opção de *rota de origem*. A idéia implícita no roteamento de origem é que ele oferece uma maneira de o transmissor impor um caminho pela interligação em redes. Para testar, por exemplo, o throughput em uma rede física especial, N, administradores do sistema podem utilizar o roteamento de origem para forçar datagramas IP a atravessar a rede N, mesmo se roteadores escolhessem normalmente um caminho que não a incluísse. A habilidade de executar esses testes é especialmente importante em um ambiente de produção, porque oferece ao gerente da rede a liberdade de rotear datagramas de usuários em redes conhecidas para operar corretamente, enquanto testem simultaneamente outras redes. É claro que esse roteamento é apenas útil para pessoas que compreendem a topologia de rede, pois o usuário médio não precisa conhecer este recurso ou utilizá-lo.

O protocolo IP aceita duas formas de roteamento de origem. Uma delas, denominada *roteamento de origem*, especifica um caminho de roteamento através da inclusão de uma seqüência de endereços IP na opção, como mostra a figura 12.

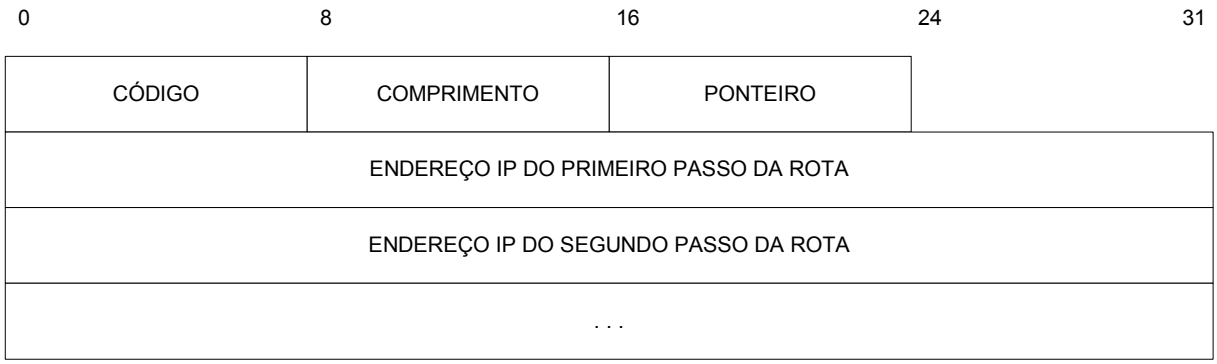

Figura 12

No roteamento restrito da origem os endereços especificam o caminho exato que o datagrama deverá seguir para chegar a seu destino. O caminho entre dois endereços consecutivos da lista deve ser composto de uma única rede física. Se um roteador não conseguir acompanhar uma rota de origem restrita, ocorrerá um erro. A outra forma, denominada *roteamento flexível de origem*, também contém uma seqüência de endereços IP. Ela determina que o datagrama deverá seguir a seqüência de endereços IP, mas permite a existência de vários passos de rota de rede entre endereços consecutivos na lista.

As duas opções de rota de origem exigem que haja roteadores ao longo do caminho para substituir itens da lista de endereços por seus endereços de rede locais. Dessa forma, quando o datagrama chega a seu destino, ele contém uma lista de todos os endereços percorridos, exatamente como a lista produzida pela opção de armazenamento de rota.

O formato de uma opção de rota de origem assemelha-se ao da opção de armazenamento de rota, mostrado anteriormente. Cada roteador analisa os campos PONTEIRO e COMPRIMENTO para verificar se a lista foi esgotada. Se for esse o caso, o ponteiro será maior que o comprimento e o roteador direcionará o datagrama até seu destino, como de costume. Se a lista não tiver sido esgotada, o roteador seguirá o ponteiro, selecionará o endereço IP, o substituirá pelo endereço do roteador e roteará o datagrama utilizando o endereço que obteve na lista.

1.17 – Opções de Indicação de Hora

A *opção de indicação de hora* funciona de forma semelhante à opção de armazenamento de rota, porque a opção de indicação de hora contém uma lista inicialmente vazia e cada roteador do caminho, da origem até o destino, preenche um item da lista. Cada entrada da lista contém dois itens de 32 bits: o endereço IP do roteador que forneceu a entrada e uma indicação do número inteiro de 32 bits que indica a hora em que o datagrama foi processado. A figura 12 mostra o formato da opção de indicação de hora.

Na figura, os campos COMPRIMENTO e PONTEIRO são usados para especificar o comprimento do espaço reservado para a opção e a localização do próximo slot não utilizado (exatamente como na opção de armazenamento de rota). O campo SOBRECARGA de quatro bits contém um contador (número inteiro) de roteadores que não poderiam fornecer uma indicação de hora, porque a opção era muito pequena.

O valor no campo FLAGS de quatro bits controla o formato exato da opção e diz como os roteadores devem fornecer indicações de hora. A tabela da figura 13 mostra os valores aceitos.

As indicações de hora fornecem a hora e a data em que um roteador trata o datagrama, expressas em milissegundos deste a meia-noite, Hora Universal. Se a representação padrão para hora não estiver disponível, o roteador poderá utilizar qualquer representação de hora local, dede que ative o bit de mais alta ordem no campo de indicação de hora. Naturalmente, as indicações de hora transmitidas por computadores autônomos nem sempre são consistentes, mesmo se representadas em hora universal; cada máquina informa a hora de acordo com seu relógio local e os relógios podem ser diferentes.

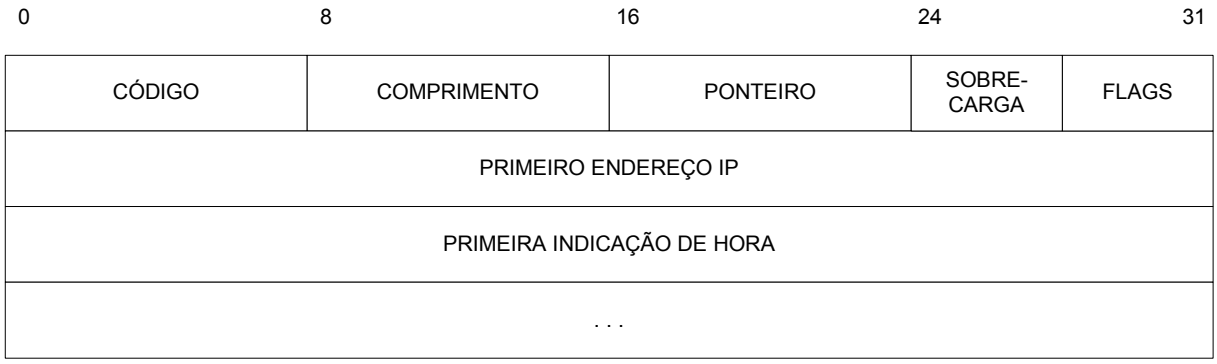

Figura 12

Pode parecer estranho que a opção de indicação de hora inclua um mecanismo para que os roteadores registrem seus endereços IP juntamente com indicações de hora, porque a opção de armazenamento de rota já oferece esse recurso. Entretanto, o registro de endereço IP com indicações de hora elimina ambigüidade. Gravar a rota juntamente com indicações de hora também é útil, porque permite que o receptor saiba exatamente o caminho seguido pelo datagrama.

Valor de Flags	Significado
0	Somente indicações de hora de registro; omitem endereços IP
1	Precede cada indicação de hora por um endereço IP (este é o formato mostrado na figura 4.17).
3	Endereços IP são especificados pelo emissor; um roteador somente registra uma indicação de hora se o próximo endereço IP da lista coincidir com o endereço IP do roteador.

Figura 13

1.18 – Como Processar Opções Durante a Fragmentação

A idéia implícita no bit CÓPIA no campo de opção CÓDIGO agora deve estar clara. Quando fragmenta um datagrama, um roteador reproduz algumas opções IP em todos os fragmentos, enquanto coloca outros em apenas um fragmento. Considere, por exemplo, a opção utilizada para registrar a rota do datagrama. Dissemos que cada fragmento deverá ser tratado como um datagrama independente; assim, não há nenhuma garantia de que todos os fragmentos sigam o mesmo caminho até o destino. Se todos os fragmentos estivessem contidos na opção de rota de registro, o destino poderia receber uma lista diferente de rotas de cada fragmento. Não poderia ser produzida uma única lista significativa de rotas para o datagrama remontado. Assim, o padrão IP especifica que a opção de armazenamento da rota dever ser copiada somente para um dos fragmentos.

Nem todas as opções IP podem ser restritas a um fragmento. Considere a opção de rota de origem, por exemplo, que especifica como um datagrama deve ser transportado através da interligação em redes. As informações de roteamento de origem precisam ser reproduzidas em todos os cabeçalhos dos

fragmentos, ou estes não seguirão a rota especificada. Assim, o campo de código para a rota de origem especifica que a opção precisa ser copiada para todos os fragmentos.