Segurança em Redes IP

Alexandre Fernandez Marques

Monografia submetida ao corpo docente do ASIT como parte dos requisitos para obter o certificado de pós-graduação

Abril 2001

Apresentação

A evolução da tecnologia da informação no curso das últimas décadas trouxe um novo significado ao termo conectividade. As possibilidades de intercâmbio de dados entre computadores, que foram exploradas desde o início de sua evolução, alcançam agora limites além das mais ambiciosas expectativas, colocando à nossa disposição recursos e funcionalidades inéditas. A notável expansão da Internet compôs um espaço virtual dinâmico e anárquico, ao alcance de milhões de pessoas ao redor do planeta, onde renovam-se a cada momento possibilidades de descoberta, entretenimento e lucro. Nesse cenário, surgiram novas formas de produção e trabalho que privilegiam a interação entre agentes e instituem a informação como bem preponderante.

É claro que essa evolução exigiu que uma interminável sequência de dificuldades técnicas fosse superada por algumas idéias brilhantes e pelo trabalho árduo de gerações de técnicos e pesquisadores. Hoje, quando encaramos informação, valor e poder como idéias convergentes, o problema que se impõe com maior urgência é a questão da segurança da informação. A ela associam-se suposições, mitos e idéias pré-fabricadas, que intrigam e divertem, mas contribuem pouco para o aperfeiçoamento de nossa compreensão sobre o assunto e muitas vezes retardam soluções e respostas.

Acredito que é responsabilidade daqueles que estão envolvidos com tarefas de gerência, criação ou pesquisa na área da Tecnologia da Informação conduzir uma discussão objetiva sobre o problema, questionando situações, revendo paradigmas e propondo alternativas. Em outras palavras, cultivando a mesma atitude que nos trouxe até aqui e que poderá fazer a evolução tecnológica seguir seu curso, no sentido de oferecer a todos serviços mais eficientes e mais seguros.

Essa é a motivação deste trabalho.

Índice

Introdução

Introdução

O problema da segurança da informação está irremediavelmente ligado à tecnologia de redes. De fato, a presença das redes de computadores em ambientes institucionais, comerciais e mesmo domésticos aproxima essa tecnologia da vida cotidiana, determinando sua participação em uma ampla gama de atividades. A existência da grande rede como espaço intermediário entre as outras redes torna toda informação potencialmente acessível a qualquer pessoa.

No cenário atual espera-se que pessoas, órgãos governamentais e instituições privadas busquem condições de beneficiar-se dos recursos oferecidos por essa tecnologia, encarando-os, conforme o caso, como objetos de interesse ou requisitos de sobrevivência.

A Internet é a referência universal em termo de redes de computadores e todas as vertentes tecnológicas parecem convergir para seus padrões. Dentre eles, destaca-se o conjunto de protocolos TCP/IP como paradigma para o intercâmbio de dados. Sua aceitação como padrão não formalizado confere à tecnologia de redes de computadores homogeneidade e consistência e condiciona hoje a própria evolução dessa tecnologia. Tenho por certo que nenhuma discussão mais profunda sobre segurança em redes de computadores pode, atualmente, deixar os protocolos TCP/IP à sua margem.

Foi assumido, desde a fase inicial desta, que o problema de segurança não admite abordagens restritivas ou perspectivas compartimentadas. Ao contrário, a complexidade de implicações e relacionamentos entre fatores e características recomenda um enfoque abrangente. Esse enfoque obriga, inúmeras vezes, à ampliação da discussão para além dos limites restritos aos aspectos de segurança do conjunto de protocolos TCP/IP. As falhas de segurança analisadas nem sempre estão diretamente relacionadas aos protocolos TCP/IP e as contramedidas estudadas são, frequentemente, aplicáveis a outros tipos de redes. Quando necessário, merecem abordagem também as particularidades e o funcionamento de componentes que interagem com o IP, assim como os efeitos que essas interações podem trazer à segurança de dados e serviços.

Este trabalho está organizado em capítulos.

No capítulo 1 são descritas as características das redes TCP/IP, focalizando seus aspectos arquiteturais e técnicos com impacto possível sobre a segurança da comunicação e dos dados. O propósito dessa descrição é estabelecer fundamentos para as análises e exposições que se seguirão, fornecendo um modelo detalhado do funcionamento do protocolo. Todavia, procura-se enfatizar nesse modelo as funções cuja relação com a segurança é imediata, o que certamente produzirá um resultado um tanto distinto do habitual.

No capítulo 2 é elaborada uma apresentação bastante resumida dos sistemas operacionais de uso mais comum em redes TCP/IP. De fato, o sistema operacional de rede empregado ou, de acordo com a moderna tendência de redes heterogêneas, a combinação de sistemas são fatores preponderantes na definição do modelo de segurança. A hegemonia de UNIX e de Windows NT como plataformas para sistemas e serviços de rede está hoje firmemente estabelecida e constitui-se ela mesma num aspecto relevante na discussão acerca da segurança de redes. Uma vez mais, busca-se destacar os traços que podem ter significado na análise da questão da segurança. Com isso, amplia-se a descrição iniciada no capítulo anterior, caracterizando o ambiente típico em que os problemas de segurança desenvolvem-se e onde devem ser tratados.

No capítulo 3 é apresentado um amplo conjunto de técnicas usadas para violação e comprometimento da segurança. Considera-se aqui que as técnicas de exploração e ataque mapeiam, numa base eminentemente prática, as vulnerabilidades dos protocolos. Portanto, a finalidade expressa dessa exposição não é outra senão a de constituir uma base para a compreensão e avaliação das soluções que serão analisadas no capítulo seguinte.

No capítulo 4, são feitas uma apresentação e uma avaliação de métodos e contramedidas que podem ser empregados com o propósito de prevenir incidentes e elevar o nível de segurança em redes que possuam as características delineadas nos capítulos 1 e 2. As contramedidas são expostas com base em seu funcionamento e aplicabilidade e avaliadas quanto à sua eficácia e limitações.

O capítulo 5 apresenta um rápido comentário sobre a elaboração e a condução de uma política de segurança que deverá apoiar e orientar procedimentos relativos à manutenção da segurança no ambiente institucional. As idéias, bastante gerais, expostas ali são livremente inspiradas na metodologia recomendada pela Módulo Security Solutions, empresa líder no mercado de segurança de informação, em treinamentos oficiais. Esta por sua vez baseia-se em padrões internacionais estabelecidos ou em desenvolvimento, como a norma ISO/IEC 17799:2000.

No capítulo 6 são estabelecidas conclusões baseadas nos aspectos discutidos nos capítulos anteriores.

Em primeira instância, este trabalho seguiu um enfoque no qual a classificação e o estabelecimento de comparações e similaridades têm importância preponderante. Contudo, as classificações propostas aqui devem ser encaradas, em sua maioria, como provisórias. Destinam-se essencialmente a facilitar a exposição de meus pontos de vista. Uma base classificatória para as análises que conduzi foi útil, a meu ver, na medida em que dá alguma coesão a um tema que tende a ramificar-se em uma infinidade de exemplos, casos particulares e aspectos correlatos.

Em segunda instância, as análises conduzidas direcionaram-se a uma avaliação de práticas, modelos e tendências. Aqui, a avaliação de produtos comerciais e soluções específicas foi deliberadamente evitada. A intenção foi a de desenvolver um pensamento orientado à observação e reconhecimento das variadas capacidades e limitações que a tecnologia de redes oferece e que afetam, de forma adversa ou não, a segurança da informação. Acredita-se que essa abordagem é capaz de fundamentar um método de julgamento e contribuir para que casos específicos possam vir a ser avaliados posteriormente, fora do escopo deste trabalho.

Capítulo 1

Arquitetura e operação do TCP/IP

Arquitetura e operação do TCP/IP

O conteúdo deste primeiro capítulo consiste numa exposição particular da forma como o conjunto de protocolos TCP/IP está estruturado e dos aspectos de seu funcionamento com efeitos relevantes sobre a segurança do ambiente. Essa exposição concentra-se portanto, em traços particulares, não devendo ser tomada como uma análise completa dos protocolos.

Os tópicos que serão desenvolvidos neste capítulo são, pela ordem, os seguintes :

- **Modelos de camadas**
- **Modelo DoD (Department of Defense)**
- **Modelo OSI (Open Systems Interconnected)**
- **Os protocolos do conjunto TCP/IP**
- **O protocolo IP (Internet Protocol)**
- **O sistema de endereçamento IP**
- **O Datagrama IP**
- **DNS – Domain Name System**
- **Roteamento**
- **Fragmentação**
- **Portas e soquetes**
- **O cabeçalho TCP**
- **A conexão TCP**
- **O cabeçalho UDP**

1. <u>Modelos de camadas</u> :

A comunicação de dados entre computadores conectados a uma rede, ou entre várias redes de computadores interconectadas entre si, é feita por meio de um complexo conjunto de tarefas ou funções, executadas por uma grande variedade de componentes de hardware e software presentes em computadores e equipamentos de conexão. A eficácia da comunicação depende em grande parte do fato de que esses

componentes, ainda que possam ser diferentes entre si, processem os dados segundo regras comuns.

Os modelos de camadas são modelos conceituais que descrevem as funções envolvidas na comunicação de dados nas redes de computadores, organizando-as em grupos denominados camadas e relacionando-as entre si numa estrutura hierárquica vertical. Embora sejam apenas construções teóricas, os modelos de camadas constituem-se numa orientação fundamental para o projeto e desenvolvimento do hardware e software de rede.

Os modelos de camadas estabelecem que na transmissão de dados cada camada recebe os dados da camada imediatamente superior, acrescenta-lhes um conjunto de informações de controle denominado cabeçalho e passa-os para a camada imediatamente inferior. Essa camada também irá acrescentar um novo cabeçalho relacionado às funções que lhe são próprias e passar o conjunto de dados recebidos mais informações de controle acrescentadas para outra camada um nível abaixo e assim, sucessivamente, até que os dados possam ser transmitidos ao meio físico de ligação da rede.

No host de destino, a camada de menor nível irá interpretar as informações que constam do cabeçalho acrescentado pela camada inferior do host de origem. Essas informações determinam a forma como os dados serão processados na camada inferior do host de destino e como essa camada irá passá-los para a camada imediatamente superior nesse host. Os dados serão agora passados de uma camada a outra, do nível mais baixo para o nível mais alto, até que os dados originais estejam disponíveis para utilização. Em cada uma das camadas as informações de controle acrescentadas pela camada correspondente no host de origem são interpretadas de modo a orientar a execução das funções próprias daquela camada.

2. O modelo DoD (Department of Defense) :

Definido no âmbito das pesquisas da ARPA (Advanced Research Projects Agency, uma agência do Departamento de Defesa norte-americano), o modelo DoD prevê a existência de quatro camadas, dispostas conforme mostra a figura a seguir :

A camada de Aplicação inclui as funções próprias das aplicações em uso na rede, tais como browsers, emuladores de terminal e utilitários. Essas funções conectam processos em execução nos hosts e estabelecem formatos compatíveis para apresentação dos dados.

A camada de Transporte, às vezes chamada da camada Host-to-Host, assegura o transporte dos dados entre transmissor e receptor, gerenciando o estabelecimento, a manutenção e o encerramento da conexão e controlando o fluxo dos dados. A camada de Transporte verifica a entrega dos pacotes e a sua ordenação no host de destino.

A camada de Internet, também referenciada como de Rede ou de Inter-redes, engloba as funções relacionadas ao roteamento dos pacotes entre as várias redes envolvidas na comunicação, como a determinação e seleção de rotas preferenciais e alternativas.

A camada Física ou de Interface de Rede e Hardware faz a interface entre os componentes de hardware de rede e os elementos lógicos que trabalham nas camadas superiores, abrangendo tarefas relativas à estrutura física da rede, às

especificações elétricas e mecânicas da mídia de transmissão e à codificação e sincronização dos bits transmitidos.

3. O modelo OSI (Open Systems Interconnected) :

Este modelo foi proposto em 1977 pelo CCITT – Comitee Consultatif International Telephonique et Telegraphique - e pela ISO – International Standards Organization – como uma abordagem alternativa ao modelo DoD, proposto pela ARPA e suas associadas.

Embora descreva as funções de rede de forma bastante semelhante ao modelo Dod, o modelo OSI é mai s detalhado, estabelecendo sete camadas, dispostas conforme mostra a figura a seguir :

As camadas de Aplicação, de Apresentação e de Sessão do modelo OSI dividem entre si as funções que no modelo DoD acham-se englobadas na camada de Aplicação. Essas funções incluem a conexão de processos e serviços entre os hosts, compressão e conversão de formato dos dados e criptografia.

A camada de Transporte desempenha basicamente as mesmas funções da camada de Transporte do modelo DoD, estabelecendo e mantendo as conexões e executando

tarefas ligadas à confirmação da entrega dos pacotes e à sua correta reordenação no host de destino.

A camada de Internet, a exemplo da camada homônima do modelo DoD, especifica métodos para a movimentação dos pacotes entre as várias redes, definindo formas de endereçamento lógico dos hosts em várias redes e determinando as rotas a serem percorridas pelos pacotes.

A camada de Link de Dados organiza os bits recebidos da camada Física em grupos lógicos, normalmente chamados frames, executando ainda funções relativas ao controle de fluxo e à identificação dos hosts dentro de um mesmo segmento de rede.

A camada Física engloba as funções relacionadas ao hardware de rede e à mídia de transmissão, de forma análoga à camada do modelo DoD.

4. Os protocolos do conjunto TCP/IP :

Protocolos são implementações de software que executam as funções necessárias à comunicação de dados conforme definidas nos modelos de camadas. A estrutura vertical dos modelos de camadas levou ao desenvolvimento de conjuntos de protocolos. Cada um dos protocolos de um conjunto executa tarefas que se relacionam, de forma mais ou menos direta, a uma camadas específica do modelo. Os conjuntos de protocolos receberam o nome de pilhas. A pilha de protocolos TCP/IP é, portanto, um conjunto composto por diversos protocolos, dentre eles o TCP (Transmission Control Protocol) e o IP (Internet Protocol).

Apresentamos a seguir uma descrição sucinta de alguns dos protocolos que compõe a pilha TCP/IP. O TCP (Transmission Control Protocol), o UDP (User Datagram Protocol) e o IP (Internet Protocol), por sua importância nas questões que esse trabalho aborda, merecerão uma análise detalhada mais adiante.

- **IP (Internet Protocol) :**

  É o protocolo responsável pelo roteamento de pacotes de dados entre as redes. Atua na camada de Internet e pode ser considerado o protocolo central do conjunto. As unidades de transferência de dados geradas pelo IP são denominadas datagramas.

- **ICMP (Internet Control Message Protocol) :**

  O ICMP é o protocolo utilizado para a transmissão de mensagens de erro no processamento de datagramas IP. As unidades de transferência de dados geradas pelo ICMP são denominadas mensagens ICMP.

- **ARP (Address Resolution Protocol) :**

  O ARP é usado para mapear os endereços IP para endereços de hardware (também chamados endereços físicos ou endereços MAC), que são utilizados para a localização de hosts no âmbito de cada rede local. Seu funcionamento consiste basicamente na conversão dos endereços IP, que constam no cabeçalho dos datagramas que chegam à rede local como endereços de hosts de destino, em endereços de hardware, o que possibilita a entrega dos datagramas ao host correto.

  Para fazer essa resolução, ARP baseia-se em consultas por meio de broadcasts e na utilização de uma tabela atualizada dinamicamente, que contém os mapeamentos já definidos a partir de consultas anteriores.
  O ARP é descrito na RFC 826.

- **RARP (Reverse Address Resolution Protocol) :**

  O RARP realiza uma resolução em sentido reverso ao daquela executada pelo ARP, ou seja, converte endereços de hardware em endereços IP. Essa função é útil em circunstâncias especiais, como por exemplo em redes onde existam

estações sem disco rígido. Essas estações conhecem seus endereços de hardware, que é registrado no firmware dos adaptadores de rede. Contudo, as estações não conhecem o endereço IP que lhes for atribuído, uma vez que não tem como armazenar essa configuração. Assim, deve ser mantida uma base de dados num servidor da rede, onde constem mapeamentos entre os endereços de hardware das estações e os endereços IP atribuídos a elas. O RARP é utilizado para consultar essa base de dados e obter o endereço IP que deve ser utilizado pela estação.

Os pacotes RARP têm estrutura muito semelhante à dos pacotes ARP, com pequenas diferenças no cabeçalho.

O RARP é descrito na RFC 903.

- **BOOTP (Bootstrap Protocol) :**

O BOOTP executa funções análogas àquelas desempenhadas pelo RARP, provendo a resolução de endereços de hardware em endereços IP a partir de uma base de dados preestabelecida e armazenada em um servidor da rede. Isso permite a inicialização de hosts com uma pilha IP reduzida, solicitando através do BOOTP o endereço IP que será utilizado e outras configurações adicionais como o endereço de gateway. O BOOTP oferece uma funcionalidade extra em relação ao RARP : a possibilidade de ter seus pacotes encaminhados através de roteadores, permitindo assim que um servidor BOOTP atenda solicitações originadas em redes distintas daquela em que está localizado.

- **Protocolos de roteamento**

Os protocolos de roteamento dão suporte às comunicações entre roteadores e outros dispositivos. Por meio de mensagens desses protocolos é feito o compartilhamento e a sincronização das informações utilizadas em tarefas de roteamento. O RIP – Routing Information Protocol – o OSPF – Open Shortest Path First – e os protocolos de roteamento externos são objeto de

comentários mais detalhados na seção deste trabalho que trata do roteamento.

- **TCP (Transmission Control Protocol) :**

  É um dos protocolos da camada de Transporte. Foi projetado para oferecer às camadas superiores serviços confiáveis e orientados à conexão. O TCP gerencia conexões ponto a ponto, isto é, com origem e destino perfeitamente definidos, e garante que os pacotes enviados tenham sido recebidos em ordem correta pelo host de destino, providenciando sua retransmissão quando requerido.

  O TCP é descrito na RFC 793.

- **UDP (User Datagram Protocol) :**

  O UDP é outro protocolo da camada de Transporte que diferencia-se do TCP pela característica de não ser orientado à conexão. Ao contrário do TCP, ele não oferece garantia de entrega de pacotes, nem recursos de recuperação de erros ou de controle de fluxo.

  O UDP funciona como uma simples interface para o protocolo IP, provendo o correto direcionamento dos datagramas entre processos em execução nos hosts de origem e destino. O mecanismo de soquetes e portas, que será descrito posteriormente, possibilita esse direcionamento.

  As funcionalidades limitadas do UDP em relação ao TCP implicam, em contrapartida, em uma sobrecarga ou overhead consideravelmente menor na comunicação. Isso recomenda seu emprego em situações nas quais a confiabilidade na entrega dos dados não é crítica, como aplicações de áudio e vídeo, e na transmissão de informações sem endereços de destino definidos.

  O UDP é descrito na RFC 768.

- **Protocolos da camada de Aplicação :**

Os protocolos dessa camada executam tarefas que compreendem desde a compatibilização de formatos de dados até o fornecimento ao usuário de comandos e interfaces.

Eles relacionam-se de maneira direta com as aplicações e são freqüentemente implementados como parte integrante delas. Isso conduz não só a uma grande diversidade de protocolos mas também à ocorrência de implementações diferentes de um mesmo protocolo. Alguns protocolos da camada de Aplicação, entretanto, obedecem a padrões bastante estáveis e têm seu uso amplamente difundido. A seguir oferecemos uma breve descrição dos mais importantes :

- **HTTP (Hypertext Transfer Protocol) :**

O HTTP é o protocolo utilizado para o acesso e apresentação de documentos em ambientes de intranets e Internet, como páginas web por exemplo. Ele define métodos de solicitação e resposta que permitem a uma aplicação cliente, geralmente um browser, conectar-se a uma aplicação servidor e executar tarefas simples de pesquisa e recuperação. O HTTP é usado também como um protocolo genérico de comunicação por agentes e gateways que utilizam protocolos diferentes de TCP/IP.

O HTTP é descrito na RFC 1945.

- **SMTP (Simple Mail Transfer Protocol) :**

O SMTP é empregado na troca de mensagens de correio eletrônico entre servidores de correio e para o envio de mensagens a partir de estações de trabalho. As funções do SMTP são limitadas ao roteamento das mensagens entre os hosts, não abrangendo a definição de interfaces locais para a

apresentação das mensagens. Para essa tarefa é comumente empregado outro protocolo, o POP3.

A RFC 821 apresenta as especificações válidas para o SMTP.

- **FTP (File Transfer Protocol) :**

As funções principais do FTP envolvem a cópia e transferência de arquivos entre hosts. Ele oferece ainda um número considerável de funcionalidades adicionais, como inspeção de diretório, autenticação do usuário, manipulação de arquivos e outras. O FTP requer que os protocolos da camada de Transporte ofereçam um serviço confiável para que ele possa estabelecer um circuito virtual entre os hosts. Tipicamente, o TCP fornece esses serviços. Em situações nas quais a confiabilidade não representa um requisito crítico, como aquelas em que a autenticação do usuário não será exigida, pode ser empregado o TFTP – Trivial File Transfer Protocol. O TFTP aceita o UDP como protocolo de Transporte.

O FTP tem suas especificações apresentadas na RFC 959 e atualizadas na RFC 2228. O TFTP é descrito na RFC 1350.

- **SNMP (Simple Network Management Protocol) :**

O SNMP é um protocolo utilizado especificamente na comunicação relacionada a tarefas de gerenciamento de redes. Essas tarefas são executadas por meio de métodos e objetos especificados na SMI – Structure and Information of Management Information – descrita na RFC 1155.

Entre esses objetos estão as NMS – Network Management Stations e os MA – Management Agents. Os MA, usualmente chamados de agentes SNMP, são executados em elementos da rede tais como hosts, servidores e gateways. Eles monitoram eventos e parâmetros da rede e armazenam o

resultado de suas observações em bases de dados chamadas MIBs - Management Information Bases. Os valores contidos nas MIBs podem ser comunicados às NMS. As NMS, freqüentemente chamadas de estações ou consoles de gerenciamento podem, por sua vez, solicitar informações aos agentes ou mesmo executar reconfigurações limitadas. A troca de informações entre as consoles de gerenciamento e os agentes é feita por meio de cinco tipos básicos de mensagens SNMP :

- **GetRequest – é uma mensagem enviada pela console de gerenciamento para recuperar um determinado valor da base de dados do agente.**

- **GetNextRequest – recupera o valor seguinte na base de dados.**

- **SetRequest – é enviada pela console a fim de alterar valores na base de dados do agente.**

- **GetResponse – é enviada pelo agente em resposta a uma das mensagens anteriores.**

- **Trap – é enviada pelo agente para notificar a console de eventos específicos, tais como a inicialização do agente, falhas de conexão etc.**

A versão dois do SNMP admite outros tipos de mensagens que estendem as funcionalidades dos tipos básicos mencionados acima.

Os agentes e as consoles são vinculados a grupos denominados comunidades SNMP. As comunidades são empregadas para organizar a troca de informações de gerenciamento e controle, garantindo-lhes acesso controlado e confidencialidade por meio de mecanismos de autenticação próprios.

O SNMP é descrito na RFC 1157.

- **MIB-II (Management Information Base Protocol) :**

  Analogamente ao SNMP, o MIB-II é também um protocolo utilizado em gerenciamento de redes.

  O MIB-II é descrito na RFC 1213.

5. **<u>O protocolo IP (Internet Protocol)</u> :**

O protocolo IP é o protocolo que atua na camada de Internet, provendo as funções necessárias ao roteamento da informação entre várias redes interconectadas. Por atuar logo acima da camada física, é o IP que define a imagem virtual ou lógica da rede.

O IP é responsável pela troca de pacotes numa base de menor esforço, ou seja, ele não oferece nenhum serviço adicional além do roteamento, exceto por alguns recursos bastante simples que permitem verificar a entrega dos pacotes e a integridade dos dados recebidos. Outras tarefas, como o estabelecimento de conexões entre os hosts envolvidos na comunicação, o controle do fluxo de dados entre eles, são deixadas para os protocolos das camadas superiores.

As primeiras especificações do protocolo IP priorizavam a necessidade estabelecida pelo DoD de manter a comunicação, mesmo na eventualidade de que algumas das redes interconectadas deixassem de operar. Com algumas das redes do sistema indisponíveis, o protocolo da camada de Internet deveria ser capaz de encontrar rapidamente rotas alternativas entre as redes remanescentes, reorientando os pacotes de uma forma tão automática quanto possível. Isso só seria possível se as funções da camada de Internet viessem a ser definidas segundo modelos não-hierárquicos, eliminando a dependência de elementos de controle que poderiam falhar, e se fosse suprimido nessa camada o emprego de serviços orientados à conexão, permitindo assim reconfigurações rápidas das rotas. Essas necesidades

foram determinantes das características atuais do protocolo e delinearam as deficiências em termos de segurança que o IP apresentaria ao ser empregado num ambiente comercial e aberto.

As especificações do protocolo IP constam das RFC 791, 919, 922 e 950, com atualizações na RFC 1349.

6. O Datagrama IP :

As unidades de transferência de dados gerenciadas pelo protocolo IP são chamadas datagramas. Um datagrama é formado pelos dados recebidos das camadas superiores, incluídas aí as informações de controle acrescentadas naquelas camadas, e pelo cabeçalho IP. O cabeçalho IP contém as informações de controle acrescentadas na camada de Internet que o protocolo IP registra ou interpreta ao executar suas funções.

O formato do cabeçalho IP é ilustrado no esquema que se segue :

<table>
<tr><td>VERS</td><td>IHL</td><td>Tipo de Serviço</td><td colspan="2">Comprimento total</td></tr>
<tr><td colspan="3">ID</td><td>FLG</td><td>Fragment offset</td></tr>
<tr><td>TTL</td><td colspan="2">Número do protocolo</td><td colspan="2">Checksum</td></tr>
<tr><td colspan="5">Endereço IP de origem</td></tr>
<tr><td colspan="5">Endereço IP de destino</td></tr>
<tr><td colspan="4">Opções IP</td><td>Padding</td></tr>
<tr><td colspan="5">Dados .....</td></tr>
</table>

Os campos que compõe o cabeçalho IP são descritos a seguir :

- **VERS – Indica a versão do protocolo IP, na maioria dos casos a versão 4, atualmente em uso.**

- **IHL – Internet Header Length – Informa o comprimento do cabeçalho IP em unidades de 32 bits.**

- **Tipo de serviço – Indica a qualidade do serviço requerida pelo datagrama. A qualidade do serviço é expressa através de um conjunto de três valores. O primeiro, denominado Prioridade, indica a precedência que deve ser dada ao datagrama. O segundo, denominado TOS, ou Type of Service, fornece informações adicionais acerca do tratamento que deve ser dado ao datagrama. O terceiro valor, denominado MBZ, tem uso exclusivamente experimental. A RFC 1349 oferece uma descrição detalhada das informações incluídas nesse campo do cabeçalho IP.**

- **Comprimento Total – Informa o comprimento total do datagrama em bytes, incluindo o cabeçalho IP e os dados.**

- **Identificação (ID) – É um número único, registrado no cabeçalho IP do datagrama pelo host de origem com o propósito de orientar a recomposição pelo host de destino de datagramas fragmentados.**

- **Flags – Este campo contém três flags cujos valores indicam atributos do datagrama relativos à fragmentação. O primeiro flag é de uso reservado e deve ser zero em todos os casos. O segundo flag, denominado DF (Don't fragment) indica quando tiver valor 1 que o datagrama não deve ser fragmentado. O terceiro flag, denominado MF (More fragments), indica quando tiver valor 0 que o datagrama não é um fragmento ou que é o último fragmento de um datagrama maior. Se tiver valor 1, indica que devem seguir-se mais fragmentos do datagrama original.**

- **Fragment offset (deslocamento do fragmento) – Este campo informa, em unidades de 8 bytes, o deslocamento do bloco de dados contido no datagrama em relação ao início do bloco de dados do datagrama original. Se o datagrama não é fragmentado ou se o datagrama é o primeiro fragmento de um**

datagrama maior, o valor indicado deverá ser zero. Os cabeçalhos dos datagramas não são computados no deslocamento.

- **TTL (Tempo de vida) – É o tempo em segundos durante o qual um datagrama permanece válido, antes de ser descartado. O valor inicial deste campo é estabelecido na criação do datagrama por um protocolo de camada superior, o TCP ou o UDP, por exemplo. Cada roteador por onde o datagrama passa subtrai um segundo do valor registrado nesse campo. Quando o valor chegar a zero o datagrama é descartado, baseado na suposição de que ele está sendo transmitido em círculos pela rede sem condições de alcançar o host de destino. Essa técnica visa evitar que datagramas permaneçam circulando indefinidamente pela rede, ocasionando tráfego desnecessário. Os roteadores são capazes de encaminhar um datagrama em menos de um segundo, mas sempre decrementam o valor do campo TTL em uma unidade a cada processamento dado ao datagrama. Embora o valor seja teoricamente expresso em segundos, na prática ele indica uma métrica de saltos de roteador a roteador, em lugar de uma métrica de tempo.**

- **Número do protocolo – Indica o protocolo de camada superior para o qual os dados contidos no datagrama devem ser passados. As indicações são feitas em valores numéricos, cada um deles correspondendo por convenção a um determinado protocolo. A STD 2 – Números Internet Designados – relaciona os números de protocolos empregados neste campo.**

- **Checksum do cabeçalho – A finalidade deste campo é permitir a verificação da integridade do cabeçalho IP a nível de bit. Contém um valor obtido a partir de uma seqüência de operações lógicas efetuadas sobre os bits que compõe o cabeçalho. Cada vez que o cabeçalho é processado o valor do checksum é recalculado. Se for obtido um valor diferente daquele que consta neste campo assume-se que pelo menos um bit do cabeçalho foi alterado indevidamente e todo o datagrama é descartado.**

- **Endereço IP de origem – Este campo contém os 32 bits que correspondem ao endereço IP do host de origem.**

- **Endereço IP de destino - Este campo contém os 32 bits que correspondem ao endereço IP do host de destino.**

- **Opções – Este campo é utilizado para a configuração de opções relacionadas com funções de controle, de indicação de erros e de teste. O campo tem tamanho variável e pode discriminar várias opções simultaneamente ou mesmo não apresentar nenhuma opção registrada.**

- **Padding – Este campo é preenchido com zeros de modo a garantir que o comprimento total do cabeçalho seja um múltiplo de 32 bits.**

7. O sistema de endereçamento IP:

No sistema de endereçamento definido pelo protocolo IP é atribuído a cada host ou a cada interface de rede um endereço único. O endereço IP é um número composto de 32 bits, divididos em quatro conjuntos de oito bits ou octetos. Uma primeira parte desses bits é usada para identificar a rede à qual o host está conectado e a parte restante é usada para identificar o host na rede.
Os endereços IP foram divididos em cinco classes, definidas pelos bits iniciais do endereço, conforme mostra a tabela a seguir :

0	**Classe A**
10	**Classe B**
110	**Classe C**
1110	**Multicast**
1111	**Reservado**

Existem ainda três faixas de endereços reservados para finalidades especiais. Esses endereços são ignorados pelas funções de roteamento executadas em roteadores e gateways, tendo assim sua utilização restrita a circunstâncias especiais. Os endereços

reservados, referenciados também como IPs falsos, internos ou privados, são os seguintes :

- **de 10.0.0.0 à 10.255.255.255 (classe A)**
- **de 172.16.0.0 à 172.31.0.0 (classe B)**
- **de 192.168.0.0 à 192.168.255.255 (classe C)**

O número de bits usados para definir a rede e o host variam conforme a classe. Exemplificando, num endereço classe A, os oito bits do primeiro octeto são utilizados para definir a rede, restando vinte e quatro bits para identificar o host. Em um endereço classe B a rede é identificada pelos bits do primeiro e do segundo octeto, num total de dezesseis bits, permanecendo outros dezesseis bits reservados para a identificação do host. Nos endereços de classe C os três primeiros octetos são usados para identificar a rede, restando apenas oito bits para a designação do host. Os endereços de classe D são utilizados para multicast e os endereços de classe E são reservados para uso futuro ou experimental.

Naturalmente, uma rede que utilize endereços de classe A irá dispor de mais bits para identificar os hosts, podendo portanto abrigar um número maior de hosts do que uma rede classe C, na qual apenas oito bits estão disponíveis para a identificação dos hosts. Note-se, entretanto, que o sistema de endereçamento permite um número maior de redes classe C, cada uma delas com um número relativamente menor de hosts, do que redes classe A, que permitem um número maior de hosts. A tabela da página seguinte ilustra esse fato, comparando o número de redes e o número de hosts por rede possíveis em cada classe.

	Redes	Hosts
Classe A	**256**	**16.777.214**
Classe B	**65536**	**65534**
Classe C	**16.777.216**	**254**

É possível notar que os endereços de classe A permitem o estabelecimento de redes com poucas sub-redes e com um número elevado de hosts, sendo dessa forma

adequados a grandes organizações. Em contrapartida, a partir de endereços classe C podem ser criadas redes apropriadas a organizações de menor porte, com um número reduzido de estações de trabalho.

A cada endereço IP é associado um segundo número, também no formato de 32 bits divididos em quatro octetos, que define quantos bits do endereço designam a rede e quantos bits designam o host. Esse número é denominado máscara de sub-rede. Ele é a base de uma técnica denominada subnetting, através da qual se consegue um nível adicional de flexibilidade na atribuição de endereços IP.

Por meio de subnetting podemos especificar para a identificação da rede um número de bits diferente dos valores padrão 8, 16 ou 24, destinando parte de um dos octetos que originalmente identificariam o host para compor a identificação da rede juntamente com os bits dos octetos precedentes. Por exemplo, um endereço classe C, que permitiria em sua configuração padrão a identificação de uma única rede com um máximo de 254 hosts pode vir a possibilitar a criação de seis sub-redes, se associado a uma máscara de sub-rede que adicione mais três bits à identificação da rede. Note-se, entretanto, que cada uma dessas redes comportará um número menor de hosts, no máximo trinta neste caso, uma vez que haverá menos bits disponíveis para identificá-los.

Os roteadores executam uma operação lógica denominada AND entre o endereço IP e a máscara de sub-rede. Em seguida, executam a mesma operação entre o segundo endereço IP e a máscara de sub-rede. Por fim, comparam bit a bit os resultados de ambas as operações. Caso os resultados sejam coincidentes, fica determinado que ambos os hosts estão localizados na mesma sub-rede.

Além dos endereços reservados, mencionados anteriormente, existem ainda casos específicos de endereçamento associados à identificação de redes e hosts. Os chamados endereços especiais são os seguintes :

- **O endereço de broadcast, representado pela atribuição de valor 1 a todos os bits que compõe a identificação de host, é usado para identificar a um**

só tempo todas as máquinas da rede, permitindo o envio de mensagens simultâneas a todas elas.

- **O endereço de rede, representado pela atribuição de valor 0 a todos os bits que compõe a identificação de host, identifica toda a rede e não uma interface específica.**

- **O endereço de loopback, representado por 127.0.0.1, identifica a própria máquina. Este endereço pode ser usado em comunicações entre processos executados na própria máquina ou em testes de configuração.**

O sistema de endereçamento IP encerra assim uma razoável flexibilidade, possibilitando a alocação de endereços para hosts e a composição de sub-redes que podem ser dimensionadas para atender à maioria das necessidades comuns.

Além disso, os endereços IP oferecem características que facilitam as funções de roteamento, possibilitando a identificação da classe do endereço pelo simples exame dos primeiros bits ou a determinação da sub-rede em que se encontra um determinado host por meio de operações lógicas de execução muito rápida.

Contudo, algumas limitações desse sistema devem ser assinaladas. A mais evidente delas é representada pela escassez de endereços. O número de endereços possíveis já não atende à demanda verificada com a expansão da Internet, que vem superando nos últimos anos todas as projeções da fase de projeto do protocolo.

A escassez de endereços IP é agravada pelo desperdício que decorre da circunstância, até certo ponto inevitável, na qual um endereço classe A é atribuído a uma organização que não irá utilizar todos os endereços de hosts permitidos. Os endereços de host vinculados àquele endereço que não forem aproveitados pela organização não poderão ser redistribuídos para outra rede, uma vez que a identificação da rede à qual pertencem é feita como parte do próprio endereço.

Outra limitação decorre da característica de que os endereços IP não são hierárquicos. Eles não se prestam a refletir as topologias organizacionais das instituições que os utilizam ou a localização geográfica das redes. Isso dificulta o suporte a técnicas de localização.

Do ponto de vista da segurança, que interessa mais diretamente a esse trabalho, nota-se que o sistema de endereçamento atualmente em uso não integra recursos de segurança próprios, relegando a outras camadas, como a camada de Aplicação, o tratamento dessa questão.

8. DNS – Domain Name System :

Domain Name System é um sistema de endereçamento de que converte nomes de host para endereços IP a partir de informações de uma base de dados. Ele é descrito pelas RFCs 1034 e 1035. O DNS consiste numa evolução de um método anterior para resolução de nomes de hosts em endereços IP. Esse método, ainda empregado opcionalmente, utiliza um arquivo texto denominado HOSTS, que continha uma lista de nomes de hosts associados a seus respectivos endereços IP.

A função mais facilmente compreendida do DNS é permitir que um host seja localizado na Internet, ou em uma rede privada de características afins, por meio de um nome mais amigável e fácil de lembrar que um endereço IP. Na verdade, outras funções ainda mais importantes são executadas pelo DNS. Discutiremos duas delas nos parágrafos seguintes.

Como dissemos, o sistema de endereçamento IP não é hierárquico e a essa característica correspondem diversas limitações. O DNS fornece um método para identificar hosts e serviços associando a cada endereço IP um ou mais nomes específicos, compostos segundo as regras de uma estrutura hierarquizada em níveis. Essa estrutura prevê a existência de domínios organizados em vários níveis a partir de uma raiz que é considerada o primeiro nível da estrutura. Os domínios podem ser associados a localizações geográficas ou a atividades específicas, possibilitando a implementação de métodos avançados de localização.

Os nomes com que o DNS designa hosts e outros recursos são denominados URLs – Uniform Resource Locators. Eles são compostos de acordo com a posição do recurso na estrutura do DNS e refletem essa localização, além de informar outras características suas, como o tipo de serviço fornecido. Um URL típico seria :

www.esquadra.mar.mil.br

Este URL identifica o serviço web, indicado por " www ", disponível no servidor denominado " esquadra ", que faz parte do domínio "mar " destinado à Marinha, que por sua vez integra um domínio de nível superior identificado por " mil " e reservado às Forças Armadas no âmbito do domínio " br ", que corresponde ao Brasil. Assim, o serviço é identificado, em meio a todos os outros existentes na Internet, a partir de um caminho composto por domínios e sub-domínios.

A base de dados usada pelo DNS para converter URLs pode conter mais de uma entrada associando URLs diferentes a um mesmo endereço IP. Isso atende a duas necessidades. Uma delas diz respeito à situação, bastante usual, em que diversos serviços são oferecidos por uma mesma máquina operando com um único endereço IP. Nesse caso, cada entrada identifica um serviço diferente, e poderíamos ter as seguintes URLs :

www.esquadra.mar.mil.br
ftp.esquadra.mar.mil.br

A outra necessidade relaciona-se ao interesse de determinadas instituições em facilitar a localização de um mesmo recurso a partir de vários URLs semelhantes, como as seguintes :

www.bancodeminas.com.br
www.bancodeminasgerais.com.br
www.bmg.com.br

Várias entradas na base de dados do DNS, cada uma delas associando um dos URL acima com o endereço IP do servidor em questão, permitiria que a solicitação de qualquer desses URLs fosse resolvida da mesma forma, fazendo com que o cliente pudesse localizar o recurso desejado.

DNS é implementado no âmbito de uma rede como um serviço disponibilizado em servidores que passam a executar, dentre outras, as seguintes tarefas de interesse :

- **manter uma base de dados com os nomes do seu domínio.**
- **responder a consultas dos clientes, resolvendo nomes de seu domínio em endereços IP.**
- **encaminhar consultas a servidores de nomes em domínios de nível superior a fim de permitir a resolução de nomes que não constam de sua base de dados.**
- **compartilhar conteúdo e atualizações da base de dados com outros servidores.**

9. <u>Roteamento</u> :

Roteamento é um conjunto de métodos e funções destinados a prover o encaminhamento correto dos datagramas entre redes interconectadas.

O roteamento é executado por dispositivos, como roteadores, comutadores e bridges, inseridos entre duas ou mais redes locais. Os dispositivos de roteamento são dotados de múltiplas interfaces de rede, cada uma delas conectada a uma rede diferente.

O problema geral do roteamento consiste em obter e manter disponíveis informações sobre a localização de cada rede ou host de destino, de modo que, ao receber um pacote, o dispositivo de roteamento seja capaz de decidir acertadamente a qual das LANs adjacentes irá encaminhá-lo, transmitindo-o pela interface de rede

correspondente. Cada itinerário capaz de viabilizar a comunicação de dados entre dois endereços conhecidos define uma rota. Na eventualidade de uma rota tornar-se indisponível, o dispositivo deve ser capaz de descobrir uma rota alternativa. Adicionalmente, na presença de várias alternativas para encaminhar o pacote, é desejável que o dispositivo possa determinar qual a rota mais econômica. Essas decisões serão tomadas confrontando o endereço de destino que consta do cabeçalho do frame ou do datagrama com as informações contidas em uma tabela, denominada tabela de roteamento, elaborada e mantida pelo dispositivo. A tabela de roteamento contém, basicamente, as seguintes informações :

- **Endereço IP ou nome da rede de destino,**
- **Máscara de sub-rede,**
- **Endereço IP ou nome associado ao gateway capaz de rotear pacotes para a rede de destino.**
- **Eventualmente, o custo em saltos ou hops para alcançar a rede de destino, estabelecido por meio de uma métrica que considera cada segmento de rede ou roteador ao longo do percurso como um hop.**

As informações de roteamento são organizadas em entradas. Um grupo de entradas típico em uma tabela de roteamento teria o seguinte formato :

Network Address	Subnet mask	Gateway
131.107.24.0	255.255.255.0	131.107.16.2
Network_1	255.255.0.0	Router A
Network_2	255.255.0.0	Router B

Quando as tabelas de roteamento são elaboradas e atualizadas manualmente, diz-se que o roteamento é estático. Quando os roteadores trocam informações entre si para subsidiar a montagem e atualização de suas tabelas de roteamento, diz-se que o roteamento é dinâmico.

O roteamento estático oferece uma simplicidade maior em redes de pequeno porte, dispensando o uso de protocolos de roteamento para suporte à comunicação entre

roteadores dinâmicos. Ele permite também que rotas preferenciais ou obrigatórias sejam definidas de acordo com as necessidades de gerenciamento da rede, evitando que elas sejam alteradas automaticamente. Contudo, sua utilização torna-se menos prática na medida em que o ambiente da rede cresce em complexidade porque o número de rotas a serem registradas manualmente pode aumentar demais. Outra limitação do roteamento estático diz respeito à necessidade de reconfiguração manual caso rotas registradas venham a se tornar inválidas pela falha de links ou roteadores.

O roteamento dinâmico, por sua vez, é mais adequado a ambientes de redes interconectadas, porque a atualização das tabelas de roteamento é realizada de forma automática, a partir de informações trocadas entre os roteadores e de métodos de manipulação dessas informações. Alterações em rotas são registradas automaticamente, reduzindo ou eliminando a necessidade de reconfiguração manual e preservando a capacidade de comunicação.

Veremos a seguir alguns dos métodos pelos quais as funções do roteamento dinâmico são implementadas.

- **<u>Roteamento de origem</u> :**

  O roteamento de origem é um método de roteamento dinâmico freqüentemente empregado em redes token ring por dispositivos denominados bridges de roteamento de origem. Na verdade, ele emprega informações de controle da camada física, situando-se assim no limite entre o roteamento e a comutação de pacotes.

  Neste trabalho iremos considerar o roteamento de origem como um método de roteamento pelas suas implicações potenciais na vulnerabilidade de certos sistemas, como será discutido mais tarde. Segundo este método, as informações de roteamento são obtidas a partir de frames recebidos dos hosts nas redes adjacentes. As informações de controle contidas nos cabeçalhos da camada física, ou cabeçalhos MAC, desses frames são utilizadas na montagem de tabelas

dinâmicas, normalmente armazenadas em RAM. As entradas da tabela são mantidas enquanto o dispositivo permanecer ativo, a menos que venham a ser redefinidas por uma nova informação sobre a rota ou sobrescritas quando a capacidade limite da tabela for atingida.

- **<u>Roteamento por vetor de distância</u> :**

O roteamento por vetor de distância caracteriza-se pela troca de informações de roteamento entre roteadores adjacentes, isto é, localizados nos limites de uma mesma LAN. A tabela de roteamento em cada roteador é montada a partir das informações recebidas dos outros roteadores. Os custos de encaminhamento são recalculados também a partir dessas informações, incrementando-se o número de hops de modo a representar o custo acrescentado pela rede adjacente. Após recalcular completamente sua tabela de roteamento, o roteador registra as rotas para cada destino conhecido com base na opção de menor custo e divulga sua nova tabela para os roteadores adjacentes.

As entradas adicionadas à tabela de roteamento de cada roteador são mantidas pelo recebimento de confirmações ou atualizações a cada comunicação com os roteadores adjacentes. Se uma entrada não for confirmada ou atualizada dentro de um determinado intervalo de tempo, ela será descartada pela suposição de haver se tornado indisponível.

Este método de roteamento envolve comunicação entre os roteadores. Essa comunicação é suportada por protocolos específicos, dos quais o mais amplamente utilizado é o RIP – Routing Information Protocol – que possui implementações para IP e IPX. Os pacotes RIP podem ser de dois tipos : solicitação e resposta. Eles são transportados em datagramas UDP/IP, normalmente em broadcast. O RIP está disponível em duas versões para uso com o Ipv4: RIP, definida na RFC 1058, e RIP-2, uma versão aperfeiçoada descrita na RFC 1723. Está definido também o RIPng, para uso exclusivo em roteadores Ipv6. A RFC 2080 descreve esse novo protocolo.

O roteamento por vetor de distância utiliza algoritmos simples e resistentes. O overhead gerado por esse método é bastante baixo, não importando em custos significativos em termos de processamento ou utilização da banda passante.

Em contrapartida o roteamento por vetor de distância apresenta falhas potenciais que limitam sua confiabilidade. O período de convergência, ou seja, o tempo necessário para que todos os roteadores divulguem e atualizem suas tabelas, representa um deles. Durante o período de convergência as funções de roteamento ficam sujeitas a falhas, dando margem à perda de pacotes. O período de convergência aumenta na razão direta do número de roteadores envolvidos, podendo ser inaceitavelmente alto em redes de grande complexidade. Em casos raros, redes com muitos roteadores podem sofrer também com problemas de consumo excessivo da banda, como conseqüência do grande volume de informação trocada entre os roteadores de vetor de distância.

Outra falha relacionada ao roteamento por vetor de distância é o problema da contagem infinita. Ele ocorre quando uma alteração inesperada na topologia da rede, ocasionada pela desativação de links ou de roteadores por exemplo, faz com que uma LAN torne-se inatingível para um determinado roteador. Caso esse roteador receba atualizações de tabela de roteamento de outro roteador adjacente das quais ainda conste a LAN inatingível, ele está sujeito a um erro de julgamento que o levará a manter a LAN como um destino válido em sua tabela de roteamento. Nesse caso, ele provavelmente registrará um novo custo de acesso àquela LAN e divulgará sua tabela para os demais roteadores. Estes, por sua vez, irão registrar a rota inválida, incrementando em um salto o custo para atingir a LAN de destino. Quando o primeiro roteador receber essas atualizações, ele irá atualizar uma vez mais sua tabela, incrementando também o custo para alcançar a rede inatingível. Esse processo ocorre tendendo ao infinito, sendo limitado na prática pelo limite definido na implementação para o incremento do custo de acesso. Normalmente o número máximo de saltos que é possível registrar é de dezesseis. Ao atingir esse valor, o destino é considerado inalcançável e o roteamento para ele falha.

A ocorrência do problema do vetor de distância pode ser controlada por algoritmos especiais que criticam a origem das atualizações recebidas, condicionando a montagem da tabela de roteamento.

- **<u>Roteamento de estado de link</u> :**

O roteamento de estado de link, às vezes chamado de estado de enlace, é outro método de roteamento, funcionando com base na comunicação entre os roteadores por meio de pacotes especiais, denominados LSP – Link State Packets. Cada roteador de estado de link gera mensagens denominadas link-state advertisements - anúncios do estado do enlace - contendo informações sobre as redes às quais o roteador está conectado. Os link-state advertisements são transmitidos em pacotes LSP para os roteadores adjacentes, que os encaminham às redes conectadas a eles, sem contudo retransmitirem para a rede de onde foram recebidos. Assim, as informações de roteamento são propagadas através de todas as redes interconectadas, permitindo a montagem de tabelas de roteamento mais completas. A seleção de rotas é feita com base nos menores custos.

Além de uma maior abrangência, as informações de roteamento usadas em estado de link tendem a oferecer maior confiabilidade do que aquelas manipuladas pelos roteadores de vetor de distância, uma vez que não são definidas de forma indireta, referindo-se em sua origem a LANs adjacentes, cujos parâmetros são bem conhecidos pelo roteador. O método apresenta um tempo de convergência menor e possibilita também o emprego de métricas de custo mais sofisticadas, nas quais podem ser considerados fatores adicionais, como a velocidade dos links. Além disso, o problema de contagem infinita não ocorre com o roteamento de estado de link.

O roteamento de estado de link utiliza protocolos específicos como o NLSP – Netware Link Services Protocol, que é usado em redes IPX, e o OSPF – Open Shortest Path First, para uso com IP.

O OSPF é definido no âmbito da RFC 2328. A RFC 1812 – Requirements for Ipv4 Routers – confere-lhe o status de único protocolo de roteamento dinâmico obrigatório. As RFCs 1245, 1246, 1247 e 1253 também fazem referência ao OSPF.

O OSPF foi projetado para suportar de maneira eficiente as funções de roteamento em ambientes de rede de médio e grande porte. Como tal, ele incorporou uma série de funcionalidades adicionais que o distinguem dos demais protocolos de roteamento. Eis algumas delas :

- **OSPF define um ambiente de roteamento organizado, estabelecendo áreas ou AS – Autonomous Systems, em função de limites geográficos ou de necessidades administrativas. No interior de um AS os roteadores compartilham e sincronizam informações de roteamento entre si. Alguns roteadores do AS são designados como roteadores de fronteira ou ASBRs – Autonomous System Border Routers. A eles cabe a tarefa de comunicar-se com roteadores em outros AS com o propósito de obter, quando necessário, informações de roteamento externas.**

- **Essa estrutura é responsável por um significativo ganho de eficiência nas tarefas de roteamento. Os roteadores OSPF no interior de um AS gerenciam um conjunto reduzido de informações de roteamento, manipulando e armazenando tabelas menores, o que diminui requisitos de hardware e custos de processamento. O tráfego necessário à troca de informações entre os roteadores e o tempo de convergência também são reduzidos.**

- **O protocolo possibilita a implementação de balanceamento de carga, distribuindo o tráfego entre as diversas rotas disponíveis para um mesmo destino.**

- **OSPF permite o uso de autenticação na troca de informações entre os roteadores.**

- **Ele é capaz também de importar rotas definidas por meio de outros protocolos, como RIP e EGP, para suas tabelas de roteamento.**

- **Máscaras de sub-rede de comprimento variável e ToS - Type of Service são suportados.**

Como afirmamos acima, alguns roteadores de um AS podem solicitar informações a roteadores localizados em AS diferentes com o propósito de determinar localização e rotas para redes remotas. Os roteadores designados para executar essa tarefa são denominados ASBRs – Autonomous System Border Routers.

A troca de informações entre ASBRs é suportada por protocolos específicos, conhecidos genericamente como protocolos de roteamento externo, em oposição aos protocolos de roteamento interno, como o RIP e o OSPF, que suportam a comunicação entre roteadores no interior de um mesmo AS. Dois dos protocolos de roteamento externo são o EGP – Exterior Gateway Protocol, definido na RFC 904 e o BGP-4 – Border Gateway Protocol, descrito pela RFC 1771, que incorpora diversos aperfeiçoamentos em relação ao EGP, reduzindo as informações necessárias ao roteamento em prol de uma maior eficiência do processo.

O roteamento constitui-se na principal função do protocolo IP. A expansão dos ambientes de redes interconectadas atribui graus de complexidade crescentes às tarefas de roteamento. Sua importância cresce conforme essa mesma tendência. A comunicação de sistemas em rede depende diretamente do roteamento. Deficiências de segurança associadas a esses dispositivos podem dar margem a comprometimentos muito amplos, afetando a um só tempo um grande número de sistemas e redes.

10. Fragmentação :

O tipo de mídia utilizado na transmissão de dados numa rede e outras características do meio físico definem um parâmetro conhecido como MTU – Maximum Transmission Unit. Ele indica o tamanho máximo do frame que pode ser transmitido através daquela rede.

Embora um datagrama IP possa ter um comprimento de até 65.535 bytes, é possível que, ao percorrer o caminho entre o host de origem e o host de destino, ele tenha de passar por redes em que a MTU seja menor que seu tamanho original.

O protocolo IP exige que todas as redes ou enlaces ao longo dos quais os datagramas serão roteados tenham uma MTU mínima de 68 bytes, valor que corresponde ao tamanho máximo do cabeçalho IP, que é de 60 bytes, acrescido de 8 bytes como comprimento mínimo dos dados. Essa exigência explica-se pela impossibilidade de dividir o cabeçalho do datagrama original sem perda de sua funcionalidade na camada. Se uma rota apresentar uma MTU abaixo de 68 bytes, o problema será tratado pela camada Física através de um processo análogo de fragmentação de frames. Se a MTU for maior que 68 bytes e menor que o comprimento total do datagrama, este deverá ser fragmentado.

A fragmentação consiste portanto na divisão de um datagrama em dois ou mais datagramas menores a fim de que eles possam ser roteados através de redes cuja MTU seja menor que seu tamanho original. Na fragmentação do datagrama são executadas as seguintes alterações :

- **O bit correspondente ao flag DF é verificado. Se estiver definido, indicando que o datagrama não deve ser fragmentado, o datagrama será descartado e uma mensagem de erro será gerada.**

- **O campo de dados do datagrama é dividido com base no valor da MTU em duas ou mais partes. É exigido que cada uma dessas partes, com exceção da última, tenha um comprimento múltiplo de 8 bytes.**

- **São gerados novos datagramas, cada um deles contendo uma parcela de dados resultante da divisão do datagrama original.**

- **O bit MF é definido em todos os novos datagrams, com exceção do que corresponde ao último fragmento**

- **O campo Fragment Offset é preenchido com o valor apropriado para cada fragmento, indicando, em unidades de 8 bytes, o deslocamento da porção de dados contida naquele fragmento em relação ao início dos dados no datagrama original.**

- **O campo Opções de cada fragmento é preenchido conforme as informações registradas nesse campo no cabeçalho do datagrama original. Algumas opções devem ser copiadas em todos os fragmentos, outras podem ser reproduzidas apenas no primeiro fragmento.**

- **O campo IHL (Comprimento do Cabeçalho) recebe o valor referente ao cabeçalho do fragmento.**

- **O campo Comprimento Total recebem o valor referente ao comprimento do fragmento.**

- **O valor do campo Checksum é recalculado.**

- **O valor do campo ID do datagrama original é mantido em todos os fragmentos, de modo a identificá-los como pertencentes a um mesmo datagrama quando forem reagrupados.**

O encaminhamento dos novos datagramas gerados pela fragmentação é feito de forma independente. Cada novo datagrama pode ser encaminhado através de rotas diferentes e podem sofrer novas fragmentações se necessário.

No host de destino os fragmentos são armazenados em um buffer específico, definido com a chegada do primeiro fragmento. O valor do campo Fragment Offset orienta a cópia dos dados na posição correta, permitindo a recomposição do datagrama original.

11. O protocolo TCP (Transmission Control Protocol) :

O TCP é o protocolo da camada de Transporte capaz de gerenciar o estabelecimento, a manutenção e o encerramento de conexões entre dois hosts, assegurando a confiabilidade da comunicação por meio de mecanismos de controle de fluxo e de recuperação de erros. Assim, quando uma aplicação comunica-se com um host remoto transmitindo e recebendo dados, o TCP gerencia o estabelecimento da conexão, ligando os processos que estão em execução no host local aos processos correspondentes no host remoto. Além disso, o TCP irá controlar o fluxo dos dados transmitidos por cada um dos hosts estabelecendo uma taxa de transmissão adequada às características de cada host e dos meios de ligação em uso. Como ainda assim existe a possibilidade de que ocorram falhas na comunicação, como a perda de pacotes ou a chegada ao destino fora da ordem em que foram transmitidos, o TCP confirmará a recepção de cada pacote transmitido, reordenando-os convenientemente no host de destino e cuidando da retransmissão dos pacotes perdidos.

As funções executadas pelo TCP simplificam o funcionamento dos protocolos da camada inferior, tipicamente o IP, permitindo que sua atuação seja limitada à entrega dos dados e conferindo-lhe assim maior velocidade e flexibilidade. Simultaneamente, o TCP dispensa os protocolos da camada superior de tarefas relacionadas à recuperação de erros ocorridos na comunicação, facilitando as implementações na camada de Aplicação e estabelecendo um método uniforme para o controle das conexões. Com a utilização do TCP como protocolo de transport, cabe às aplicações apenas fornecer um fluxo contínuo de bytes a serem transmitidos.

O TCP é descrito na RFC 793.

12. Portas e Soquetes :

A comunicação entre computadores em rede freqüentemente estabelece casos nos quais uma única aplicação em execução num host precisa atender a conexões simultâneas de vários clientes. Um servidor web exemplifica bem essa situação. Existe ainda a possibilidade de que cada um de vários processos em execução num host deva comunicar-se com um determinado processo em outro host. As diferenças verificadas na forma como sistemas operacionais distintos identificam os processos ampliam essas dificuldades. O mecanismo de portas e soquetes oferece uma solução adequada ao problema.

As portas são números de 16 bits que permitem ao protocolo de Transporte identificar um processo, que pode ser na prática uma aplicação ou outro protocolo de nível superior, como o destinatário correto dos dados que devem ser entregues. Processos padronizados identificam-se por números na faixa de 1 a 1023. Essas portas são chamadas por isso de portas bem conhecidas. As portas de 1024 a 65535 são denominadas portas efêmeras e destinam-se a identificar processos não padronizados.

A designação das portas efêmeras é feita de maneira dinâmica pelo TCP/IP. O processo que inicia a comunicação solicita ao conjunto uma porta disponível e o valor fornecido é informado no cabeçalho dos segmentos TCP ou dos datagramas UDP que são enviados.

Os soquetes constituem uma interface de programação genérica para comunicação. Eles foram introduzidos com o UNIX BSD 4.2 e tiveram ampla aceitação. Um soquete define um endereço de soquete, às vezes chamado de endereço de transporte ou simplesmente de soquete, do qual fazem parte três elementos:

- **O protocolo de transporte utilizado.**
- **O endereço do host local.**
- **O processo em comunicação, identificado por uma porta.**

Um exemplo de soquete para uma conexão sobre TCP/IP seria :

tcp, 200.244.203.37, 80

Quando uma conversação entre dois hosts é estabelecida, cria-se uma associação entre os soquetes de cada host. Como ambos os hosts devem utilizar o mesmo protocolo de transporte, a associação inclui cinco elementos:

- **O protocolo de transporte utilizado.**
- **O endereço do host local.**
- **O processo em comunicação no host local, identificado por uma porta.**
- **O endereço do host remoto.**
- **O processo em comunicação no host remoto, identificado por uma porta.**

Um exemplo de associação para uma conexão sobre TCP/IP seria :

tcp, 200.244.203.37, 80, 200.244.203.65, 2016,

Dessa forma, uma conexão lógica entre dois processos em hosts distintos pode ser perfeitamente identificada, a partir dos soquetes que correspondem aos aplicativos que se comunicam de cada lado da conexão. Se houver a necessidade de estabelecerem-se conexões simultâneas entre os dois hosts, o mecanismo de portas e soquetes dará suporte às múltiplas conexões, identificando os processos envolvidos e permitindo seu controle pelos protocolos da camada de Transporte. A comunicação interprocessual entre os hosts é feita por meios de várias conexões únicas, conferindo às conversações sobre TCP/IP características de multiplexação.

13. O cabeçalho TCP :

As unidades de transferência de dados gerenciadas pelo protocolo TCP são chamadas segmentos. O cabeçalho TCP contém informações necessárias ao gerenciamento das conexões lógicas controladas pelo protocolo, incluindo-se as relacionadas ao controle do fluxo de dados transmitidos, à reordenação dos dados no destino e à recuperação de erros. A forma pela qual essas informações são utilizadas será objeto de uma exposição detalhada na descrição do mecanismo de estabelecimento da conexão TCP, que se segue à descrição dos campos do cabeçalho.

O formato do cabeçalho TCP é ilustrado no esquema que se segue :

<table>
<tr><td colspan="4">Porta de Origem</td><td colspan="5">Porta de Destino</td></tr>
<tr><td colspan="9">Número Seqüencial</td></tr>
<tr><td colspan="9">Número de Reconhecimento</td></tr>
<tr><td colspan="2">Deslocamento dos dados</td><td colspan="2">Reservado</td><td>URG</td><td>ACK</td><td>PSH</td><td>RST</td><td>SYN</td><td>FIN</td><td>Janela</td></tr>
<tr><td colspan="4">Checksum</td><td colspan="5">Ponteiro Urgente</td></tr>
<tr><td colspan="4">Opções</td><td colspan="5">Preenchimento</td></tr>
<tr><td colspan="9">Dados .....</td></tr>
</table>

- **Porta de Origem – Neste campo é registrado o número da porta no host de origem que será usada na conexão.**

- **Porta de Destino - – Neste campo é registrado o número da porta no host de destino que será usada na conexão.**

- **Sequence Number – Indica o número seqüencial atribuído ao primeiro byte dos dados enviados neste segmento. No caso de o segmento ser o primeiro de uma conexão, o sequence number é determinado por um processo que envolve variáveis aleatórias e é denominado ISN - Initial Sequence Number. O sequence**

number atribuído ao primeiro byte de dados deste segmento será igual ao valor do ISN mais um.

- **Número de Reconhecimento – Indica o número seqüencial que o receptor espera receber no próximo segmento.**

- **Deslocamento dos Dados – Este campo indica a posição no segmento onde acaba o cabeçalho TCP e começam os dados, informando o número de palavras ou grupos de 32 bits contidos no cabeçalho.**

- **Reservado – Este campo acha-se reservado para uso futuro. Os seis bits que o compõe devem ser configurados com valor zero.**

- **Bits de Controle – Este campo é composto por seis bits. Conforme o valor atribuído a eles, os bits de controle sinalizam eventos relacionados à abertura da conexão, à sincronização dos dados transmitidos e recebidos e ao encerramento da conexão. É exposta, a seguir, a função de cada um deles :**

- **URG – Quando este bit tiver valor 1 indicará que o conteúdo do campo Ponteiro Urgente deve ser considerado. Se valer zero, o Ponteiro Urgente será ignorado.**

- **ACK - Quando este bit estiver definido indicará que o Número de Reconhecimento é significativo. Caso valha zero, o conteúdo do campo não será considerado.**

- **PSH – Quando estiver configurado com valor 1 iniciará uma função de descarga (push).**

- **RST – Este bit, quando definido, provocará a reinicialização da conexão.**

- **SYN – Quando estiver definido este bit irá forçar a sincronização dos números seqüenciais, tipicamente no início de uma conexão.**

- **FIN – Indica, quando definido, que não há mais dados a serem transmitidos pelo host de origem, encerrando a conexão.**

- **Janela – Informa, nos segmentos ACK, o número de bytes que o host está em condições de receber.**

- **Checksum – Este campo contém um valor usado para verificar a integridade do cabeçalho TCP e detectar falhas na comunicação. O valor é obtido a partir de uma operação de soma em complemento dos valores dos campos do cabeçalho TCP, dos dados do segmento e dos campos que compõe o chamado pseudo-cabeçalho. O pseudo-cabeçalho é composto pelos seguintes campos : Endereço IP de Origem, Endereço IP de Destino, Número do Protocolo e Comprimento do Segmento TCP. O pseudo-cabeçalho tem a única finalidade de servir ao cálculo do checksum, não sendo registrado ou transmitido com o segmento. Contudo, ele torna a verificação mais ampla incluindo no cálculo valores do datagrama IP.**
  Os valores dos campos checksum e padding do cabeçalho TCP não são considerados para o cálculo do checksum.

- **Ponteiro urgente – Este campo, considerado apenas quando o bit de controle URG estiver definido, informa o número seqüencial do primeiro byte que se segue a um conjunto de dados classificados como urgentes.**

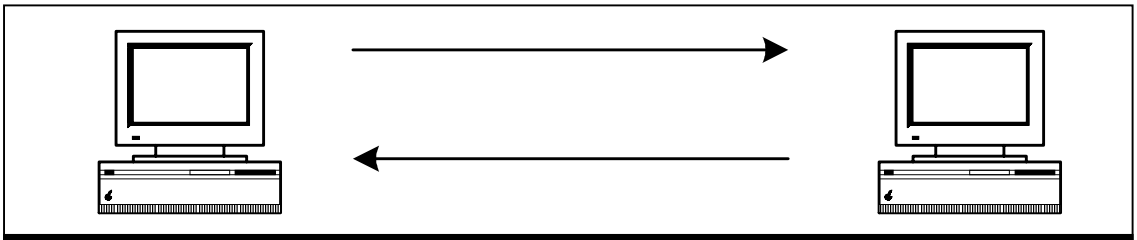
- **Opções – De forma similar ao campo Opções do cabeçalho IP, este campo apresenta um conjunto de valores que indicam opções não obrigatórias, relativas ao controle de fluxo e recuperação de erros.**

- **Padding - Este campo é preenchido com zeros de modo a garantir que o comprimento total do cabeçalho seja um múltiplo de 32 bits.**

14. A conexão TCP :

O TCP estabelece conexões unidirecionais e ponto a ponto, ou seja, a origem e o destino da conexão devem ser perfeitamente definidos e os dados seguem sempre na

mesma direção. A condição de host de origem não pode ser acumulada, na mesma conexão, com a condição de host de destino. Quando dois hosts comunicam-se através de uma conexão TCP transmitindo e recebendo dados, são estabelecidas na verdade duas conexões, uma em cada sentido. Através de uma delas um host transmite dados e é, portanto, o host de origem. Através da outra o mesmo host recebe dados, representando então o host de destino. Como as duas conexões são operadas simultaneamente, estabelece-se um canal full-duplex, em que a comunicação é feita ao mesmo tempo em dois sentidos, como ilustrado abaixo :

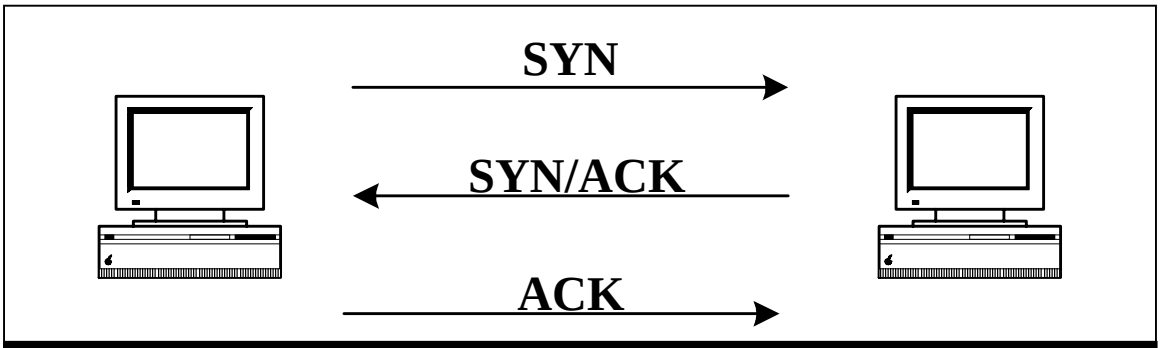


Conexão TCP full – duplex.

A conexão TCP é estabelecida por meio de um processo conhecido como three-way handshake. Quando uma aplicação, por meio de uma API como WinSock, por exemplo, solicita comunicação com um host remoto é gerado um segmento TCP com o bit SYN ligado em seu cabeçalho. Esse segmento, que passaremos a chamar simplesmente de SYN, contém ainda um valor inicial para o sequence number e o tamanho atual da janela. Essas informações de controle terão seu uso esclarecido mais adiante.

Ao receber o segmento SYN, o host de destino envia um segmento com o propósito de participar o reconhecimento do pedido de conexão. O cabeçalho desse segmento tem os bits SYN e ACK definidos, pelo que passaremos a nos referir a ele como SYN/ACK. O host de origem recebe o segmento SYN/ACK e responde com um segmento em cujo cabeçalho o bit ACK está definido.

Este terceiro passo conclui o estabelecimento da conexão, ficando ambos os hosts em condições de intercambiarem dados por meio dela. As três etapas desse processo são esquematizadas na ilustração a seguir :



Estabelecimento conexão TCP (three-way handshake).

No decorrer do processo de comunicação dos hosts, os protocolos da camada de Aplicação irão fornecer ao TCP um fluxo contínuo de bytes a serem transmitidos, sem qualquer divisão ou formatação específica para a camada de Transporte. É responsabilidade do TCP dividir esse fluxo de dados em segmentos de tamanho adequado à comunicação, considerando a capacidade do host de destino para recebê-los sem erros. Na verdade, no host de destino os dados recebidos da rede são armazenados em um buffer para reordenamento e verificação antes de serem passados à camada de Aplicação. A capacidade desse buffer, associada a outros fatores como congestionamentos em redes intermediárias ou erros na ordem em que os segmentos são recebidos, irá determinar um volume adequado de dados que o host de destino pode receber. Esse valor será estabelecido como o tamanho da janela ou seja, a quantidade de bytes que o host de origem pode enviar de uma só vez sem criar um fluxo excessivo de dados que leve a falhas de comunicação. O tamanho da janela é informado pelo host de destino quando do estabelecimento da conexão, no cabeçalho do segmento SYN/ACK.

Os bytes que precisam ser transmitidos pelo host de origem são armazenados também num buffer de saída, e a janela indica quantos desses bytes podem ser transmitidos pela conexão nas condições atuais. O host de origem pode agrupar todos os bytes contidos na janela em segmentos e transmiti-los antes que o host de destino confirme o seu recebimento, mas somente poderá transmitir os bytes seguintes após obter essa confirmação.

Enquanto permanecem ordenados no buffer de saída aguardando a transmissão, cada um dos bytes recebe um sequence number, a partir do número consecutivo (n+1) ao sequence number inicial, que foi informado pelo host de origem no cabeçalho do segmento SYN ao estabelecer-se a conexão. Como cada byte recebe um sequence number, mas os dados são transmitidos em segmentos que reúnem vários bytes, o sequence number informado no cabeçalho TCP é o sequence number atribuído ao primeiro dos bytes de dados contidos no segmento. À medida em que os segmentos vão sendo transmitidos, eles permanecem armazenados no buffer e é ativado para cada um deles um contador de tempo de retransmissão no host de origem. Quando esse contador chega a zero o segmento é retransmitido.

Quando o host de destino recebe um determinado número de segmentos, usualmente dois, ele envia um segmento ACK no cabeçalho do qual informa o como número de reconhecimento o sequence number do próximo segmento que espera receber, considerando o sequence number do último segmento recebido e o tamanho dos segmentos. Exemplificando, se o sequence number do último segmento recebido foi 2048 e o tamanho dos segmentos é de 512 bytes, o valor informado no campo número de reconhecimento do segmento ACK será 2560. Caso o host de destino não receba um dos segmentos transmitidos, ele será capaz de detectar o erro pelo exame do sequence number dos segmentos que recebeu e continuará informando o sequence number do segmento que falta no campo número de reconhecimento dos segmentos ACK que enviar ao host de origem, até que venha a recebê-lo.

Quando o host de origem recebe um ACK referente ao primeiro dos segmentos incluídos no buffer de saída ele "desliza a janela" para adiante, de modo a incluir o próximo segmento, que poderá agora ser transmitido. Ao fim da comunicação, o fechamento da conexão é executado de forma similar ao estabelecimento da mesma. Um segmento com o bit FIN definido em seu cabeçalho é enviado pelo host de origem e a conexão é encerrada.

15. O cabeçalho UDP :

As unidades de transferência de dados gerenciadas pelo protocolo UDP são chamadas datagramas. Uma vez que o protocolo não estabelece conexões como o TCP, funcionando apenas como uma interface para o protocolo IP, seu cabeçalho é consideravelmente mais simples, o que contribui para reduzir a sobrecarga que as informações de controle trazem à comunicação. Em contrapartida, o UDP não é capaz de oferecer a confiabilidade e os recursos de controle de fluxo e recuperação de erros obtidos por meio da utilização do TCP. O UDP limita-se a direcionar, por meio do mecanismo de soquetes e portas, os datagramas trocados entre processos nos hosts de origem e destino.

O formato do cabeçalho UDP é apresentado a seguir :

Porta de Origem	**Porta de Destino**
Comprimento	**Checksum**
Dados	

- **Porta de Origem – Neste campo é registrado o número da porta no host de origem que será usada na conexão.**

- **Porta de Destino - Neste campo é registrado o número da porta no host de destino que será usada na conexão.**

- **Comprimento - Este campo informa o comprimento total do datagrama em bytes, incluindo o cabeçalho.**

- **Checksum – Da mesma forma que no cabeçalho TCP, este campo é utilizado para verificação de integridade. O valor registrado aqui é obtido pelo mesmo método empregado no TCP, inclusive com o uso dos valores dos campos do pseudo-cabeçalho.**

Capítulo 2

Sistemas Operacionais de Rede

Sistemas Operacionais de Rede

Neste capítulo são descritos os sistemas operacionais de emprego mais freqüente em redes IP. Uma caracterização completa de cada um deles está além de nossos propósitos. Sua descrição é traçada em termos de aspectos gerais e focalizada no que possa ter implicações sobre a segurança de dados e serviços, seja representando capacidades, seja condicionando vulnerabilidades. Cabe lembrar que os roteadores e outros dispositivos de rede executam, por vezes, sistemas operacionais proprietários, que podem ser sensíveis às mesmas vulnerabilidades que os sistemas aqui descritos ou ainda a vulnerabilidades específicas. Pela particularidade representada por esses casos, o presente trabalho não se ocupa da sua descrição.

Serão descritos os seguintes sistemas operacionais de rede :

- **UNIX e LINUX**
- **Microsoft Windows NT**
- **Novell Netware**

1. UNIX e LINUX :

O UNIX tem sua origem mais remota no Multics, um sistema operacional desenvolvido no âmbito de um projeto conjunto da Bell Labs, subsidiária da AT&T, do Massachussets Institute of Technology e da General Electric. Ao final da década de sessenta, o projeto sofreu uma descontinuidade. Movido pela necessidade de continuar utilizando aplicações criadas para o Multics, Ken Thompson, um dos programadores envolvidos no projeto, deu início ao desenvolvimento de um sistema operacional de características semelhantes. Como as modestas ambições de Ken Thompson apontavam, então, para um sistema mono-usuário, o sistema, surgido dessa iniciativa pessoal foi chamado de UNIX, por oposição ao Multics.

A capacidade de suporte a múltiplos usuários foi um dos primeiros aperfeiçoamentos incorporados ao UNIX. Em 1973 o sistema foi recodificado em C. Essa etapa

representou um momento extremamente importante no desenvolvimento do sistema. C era uma linguagem de alto nível em evolução e seu aperfeiçoamento em paralelo com a recodificação do UNIX durante esse período veio trazer grande consistência e flexibilidade a ambos.

Versões do UNIX e da linguagem C foram oferecidos à comunidade acadêmica para uso e pesquisa. Isso acelerou a evolução do sistema operacional e provocou o surgimento de versões trazidas ao mercado por empresas como a IBM, que criou o AIX, a Hewlett-Packard, que desenvolveu o HP-UX e a Sun, com o Solaris. Curiosamente, o UNIX, concebido como um sistema mono-usuário e de uso restrito, assumia uma tendência de diversificação contínua que resistiria a todas as iniciativas de padronização.

O LINUX pode ser considerado uma das versões de UNIX mais recentes e revolucionárias. Diretamente associado aos conceitos de software aberto e gratuito, seu lançamento obteve grande repercussão, num momento em que o monopólio da Microsoft no negócio de sistemas operacionais enfrenta contestações generalizadas. Em 1991, Linus Torvalds, então estudante da Universidade de Helsinki, deu início ao desenvolvimento de um sistema operacional baseado no UNIX. À sua iniciativa pessoal juntaram-se colaboradores de todo o mundo, num esforço cooperativo conduzido majoritariamente por meio da Internet. O sistema beneficia-se de características herdadas do UNIX e da condição de ser gratuito para expandir seu emprego em ambientes diversos, como empresas, instituições educacionais e entre usuários domésticos. Assim como aconteceu com o UNIX, o suporte eficiente a serviços de rede tem levado o LINUX a ser adotado como plataforma para servidores web, roteadores baseados em software e outras aplicações.

As diferenças entre os dialetos UNIX e as distribuições LINUX podem, eventualmente, tornar menos precisa a apresentação feita aqui. No entanto, a bem da objetividade, fiz opção por não explorar essas diferenças e descrever os sistemas com base em suas características comuns.

Algumas das principais características compartilhadas pelo UNIX e pelo LINUX são descritas a seguir :

- **<u>Portabilidade</u> :**

  UNIX é um sistema aberto e flexível. Suas características conferem-lhe grande capacidade de adaptação, podendo ser executado sobre plataformas de hardware tão diversas quanto Intel, RISC, SPARC, Alpha e outras.

- **<u>Arquitetura modular e flexível</u> :**

  A arquitetura do UNIX e do LINUX apresenta um núcleo do sistema operacional, denominado kernel, ao qual podem ser associados módulos e aplicativos em uma variedade praticamente ilimitada. O kernel é responsável pela definição e controle de um ambiente básico de execução para processos e tarefas, gerenciando as funções de processamento e as interações entre software e hardware.

  Uma segunda camada na arquitetura do sistema é representada pelo shell. O shell executa funções de interpretação de comandos, permitindo a interação dos usuários com o software instalado. O shell de UNIX/LINUX incorpora uma linguagem de programação básica, por meio da qual podem ser gerados scripts para simplificação de comandos e automatização de tarefas. Há duas versões de shell usadas na maioria das versões e distribuições de UNIX/LINUX, o Bourne Shell, desenvolvido pela Bell Labs e o C Shell, originado na Berkeley University. As versões apresentam pequenas diferenças entre os comandos disponíveis em uma e outra e oferecem funcionalidades bastante similares.

  A arquitetura do UNIX/LINUX completa-se com uma terceira camada, à qual pertencem os aplicativos e utilitários.

  O acesso ao código do sistema permite que ambos os sistemas sejam adaptados a necessidades específicas de cada forma de utilização. No caso do LINUX é

possível recompilar o sistema de modo a adequá-lo a determinado ambiente ou a uma função particular.

A flexibilidade oferecida por UNIX/LINUX constitui-se, com toda certeza, em um dos maiores atrativos dos sistemas, possibilitando-lhe oferecer funcionalidades bastante diversas e permitindo seu emprego em uma ampla gama de funções e tarefas. Do ponto de vista da manutenção da segurança, todavia, ela motiva grandes preocupações. O controle sobre configurações e recursos deve ser exercido com elevado grau de detalhamento. A investigação de falhas e comprometimentos deve ser contínua e abrangente, incluindo a própria estrutura do sistema, os arquivos de configuração e as alterações em módulos e aplicativos.

- **<u>Versões e dialetos</u> :**

Como foi mencionado anteriormente, o desenvolvimento de versões proprietárias do UNIX levou à existência de um grande número de implementações que, embora compartilhem muitas características comuns, guardam entre si diferenças importantes e revelam, com freqüência, limitações à interoperabilidade.

De uma forma semelhante, o LINUX é apresentado em pacotes organizados que contém, além do núcleo do sistema, conjuntos específicos de aplicações, interfaces e módulos de apoio. Esses pacotes, usualmente chamados de distribuições, apresentam também características que os diferenciam um dos outros, constituindo-se dessa forma em similares das versões comerciais do UNIX.

A existência desses dialetos ou "sabores" do UNIX, como são chamados, e das várias distribuições do LINUX, torna as tarefas de administração desses sistemas bastante complexas e especializadas. Métodos específicos de segurança podem ser adequados a uma versão e revelar-se ineficientes em outra. Diferentemente do que ocorre com sistemas operacionais padronizados, acompanhar a descoberta de novas vulnerabilidades em um sistema aberto como UNIX/LINUX, verificar sua ocorrência nos hosts sob nossa responsabilidade e treinar pessoal de apoio e

usuários para corrigir ou evitar essas falhas envolve um esforço contínuo e exige alto nível de qualificação. Os custos desse processo tendem a ser mais altos do que o verificado em sistemas padronizados e precisam ser considerados em tempo de planejamento e organização.

- **Suporte avançado a serviços de rede :**

O UNIX oferece a possibilidade de execução de processos em ambiente de multitarefa, ou seja, é capaz de gerenciar a execução simultânea de várias tarefas por um mesmo processador. O suporte a ambiente multitarefa e multiusuário é também incorporado ao kernel do LINUX. Em ambos os sistemas esse suporte estende-se à possibilidade de que cada uma das tarefas em execução seja controlada por um usuário diferente. Essa característica tem recomendado a utilização do UNIX/LINUX em servidores de grande porte, em detrimento de outros sistemas operacionais de rede que oferecem a mesma capacidade em grau mais limitado, como o Windows NT.

Como decorrência disso, serviços de rede são integrados ao sistema operacional de forma estreita e eficiente. UNIX/LINUX oferece a possibilidade de que um usuário remoto conecte-se ao sistema e a partir daí seja capaz de executar tarefas e acessar recursos como se estivesse operando localmente. Diz-se nesse caso que o usuário remoto utiliza um shell do sistema em questão. O enorme ganho funcional representado por essa possibilidade implica em condições de segurança menos restritivas. Isso explica a abordagem de McClure e Scambray segundo a qual um usuário somente deve ser considerado como usuário remoto enquanto não tem acesso a um shell do sistema. Esse usuário pode ser considerado local a partir do momento em que obteve um shell, tendo à sua disposição o mesmo conjunto de comandos e interfaces que usaria se tivesse acesso físico à máquina. A obtenção de um shell é um passo intermediário importante na maioria das técnicas de ataque orientadas a UNIX/LINUX.

- **<u>Contas de usuário</u> :**

Qualquer ação ou tarefa executada em uma máquina UNIX/LINUX, mesmo uma tarefa executada localmente, é dependente de uma conta de usuário. Isso coloca as contas de usuário numa posição central no modelo de segurança desses sistemas. Nomes de usuário com contas válidas e senhas são informações críticas em qualquer sistema operacional, mas em UNIX/LINUX representam um alvo prioritário para atacantes e invasores. A segurança de login deve ser, portanto, objeto da maior atenção por parte dos administradores, uma vez que seu comprometimento pode levar a danos de grande monta.

As contas de usuário são criadas como entradas no arquivo /etc/passwd. Cada entrada discrimina o nome do usuário e a senha a ser utilizada em seu login, além de informações sobre participação em grupos e diretório padrão do usuário. O arquivo é composto em modo texto e pode ser aberto ou editado por qualquer editor de textos comum. Além disso, ele deve ser legível para qualquer usuário. Evidentemente isso representa uma vulnerabilidade importante, ainda que as senhas sejam registradas em /etc/passwd sob forma criptografada. Para controlar essa vulnerabilidade é adotada por quase todas as versões e distribuições de UNIX/LINUX a técnica de sombreamento (shadowing). O sombreamento consiste na substituição das entradas no arquivo /etc/passwd por outras entradas que contém apenas índices para entradas reais contidas em outro arquivo, /etc/shadow, protegido por restrições de acesso.

Quando da sua criação, cada conta de usuário é associada a um grupo. Embora seja possível indicar manualmente a que grupo o usuário que está sendo criado deve ser adicionado, existem procedimentos padrão que são seguidos na falta de indicações específicas. Em algumas versões é criado um grupo exclusivo para cada usuário acrescentado ao sistema. Em outras, todos os usuários são incluídos em um grupo padrão.

Todas as versões e distribuições de UNIX/LINUX utilizam uma conta padrão com direitos irrestritos sobre o sistema. Essa conta, denominada Root, é criada

no processo de instalação e confere a quem a utiliza a possibilidade de executar qualquer ação, incluindo a criação de usuários e a alteração de direitos e permissões.

UNIX/LINUX, em suas várias versões ou distribuições, apresentam contas de usuário padrão, que utilizam senhas default ou, em alguns casos, não exigem sequer o uso de senha para conectar-se ao sistema. Algumas dessas contas são criadas na instalação do sistema operacional. Outras são criadas pela instalação de pacotes de aplicativos. Em qualquer caso, é importante adotar medidas de controle como desabilitar as contas desnecessárias ou alterar as senhas padrão, a fim de evitar que invasores aproveitem-se dessa facilidade para obter acesso a dados ou, o que é mais freqüente, a informações sobre o sistema.

Outro aspecto a considerar é a utilização dos usuários confiáveis. Essa prática simplifica o processo de autenticação de usuários, exigindo apenas um login para que o usuário possa ter acesso a todos os hosts de uma rede que tiverem estabelecido relações de confiança com o host onde ele está operando. As implicações desse método sobre a segurança da rede são evidentes e exigem um estrito controle sobre as relações de confiança configuradas.

- **NFS - Network File System :**

O NFS é um sistema de arquivos desenvolvido pela Sun Microsystems, Inc., que permite aos usuários de estações de trabalho acessar arquivos e diretórios remotos de uma rede como se fossem arquivos e diretórios locais. Embora não seja o único sistema de arquivos usado em UNIX/LINUX, o NFS está presente na maioria dos dialetos e distribuições.

O NFS define uma estrutura hierárquica de diretórios que inclui vários componentes padrão. A estrutura tem seu nível mais elevado na raiz ou root, representado por /. A organização do sistema de arquivos é registrada em um arquivo especial denominado Diretório. O esquema a seguir ilustra uma estrutura típica de sistema de arquivos em UNIX.

/ ..... root - raiz do sistema de arquivos

/kernel ..... kernel – núcleo do sistema operacional

/bin ..... arquivos executáveis

/dev ..... dispositivos

/etc ..... arquivos de configuração

/usr ..... aplicativos

/usr/bin ..... outros executáveis

/tmp ..... arquivos temporários

/var ..... arquivos de log e outros

/opt ..... pacotes de software

/export/home ..... diretórios de usuários

O acesso a recursos do sistema é baseado no acesso aos arquivos executáveis ou aos arquivos de configuração correspondentes a esses recursos. Em termos práticos, para utilizar um recurso, um usuário precisará ter o direito de executar programas ou aplicativos associados a esse recurso. Para alterar configurações desse recurso, o usuário precisará ter o direito de editar o arquivo de configuração que define suas formas de utilização. Dessa forma, em UNIX/LINUX as restrições que podem ser estabelecidas sobre recursos do sistema relacionam-se diretamente aos métodos de controle de acesso a arquivos.

UNIX/LINUX controla o acesso a arquivos com base em um conjunto de permissões associadas a cada arquivo.

Há três tipos de permissão aplicáveis a arquivos :

r – permite a leitura do arquivo ou diretório.

w – permite a gravação ou alteração do arquivo ou diretório.

x – permite a execução de um arquivo ou o exame de conteúdo de um diretório.

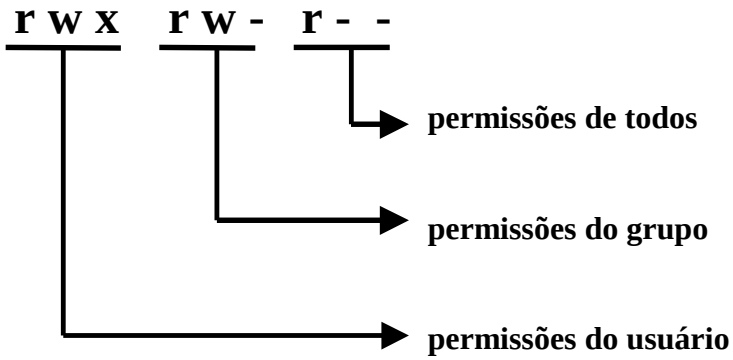
Essas permissões podem ser associadas a três tipos de usuários :

u – indica o usuário definido como proprietário do arquivo ou diretório.

g – indica usuários membros do grupo ao qual o proprietário do arquivo ou diretório pertence.

o – indica todos os demais usuários do sistema.

Assim, as permissões de acesso são expressas sob a forma de uma string onde cada seqüência de três caracteres refere-se a um tipo de usuário ou grupo, na ordem em que foram apresentados acima. As permissões não concedidas são representadas por um traço na posição correspondente. Exemplificando, uma string como a representada abaixo indica que o proprietário do arquivo tem direitos de leitura, gravação e execução sobre o arquivo. O grupo ao qual o usuário está vinculado tem direito de leitura e gravação e os demais usuários têm apenas o direito de leitura.



As permissões de acesso podem também ser expressas sob a forma de um número de três algarismos, composto de acordo com a seguinte convenção :

0 = Sem acesso

1 = Permissão de execução

2 = Permissão de gravação

3 = Permissão de gravação e execução

4 = Permissão de leitura

5 = Permissão de leitura e execução

6 = Permissão de leitura e gravação

7 = Permissão de leitura, gravação e execução

Assim, um arquivo marcado com o número 754 tem associado a si as seguintes permissões :

- **o proprietário do arquivo pode ler, alterar e executar o arquivo (7)**
- **o grupo pode ler e executar o arquivo (5)**
- **todos os usuários podem apenas ler o arquivo (4)**

Além das permissões descritas acima, existem duas permissões especiais, conhecidas como SUID e SGID. Essas permissões fazem com que um arquivo sempre seja executado com os direitos do proprietário ou do grupo, respectivamente, mesmo que o usuário que o executa não tenha privilégios especiais. Um arquivo com SUID de root será executado por qualquer usuário como se este fosse root.. É fácil perceber que as permissões SUID e SGID devem ser usadas com cuidado para evitar sua exploração em prejuízo da segurança.

Além dos fatos que se referem a permissões de arquivos, o NFS está associado a diversas vulnerabilidades próprias, que incluem desde condições de buffer overflow a partir de comandos específicos até a possibilidade de exportar total ou parcialmente o sistema de arquivos de um host UNIX para um host remoto. Esse resultado permitiria ao atacante examinar o conteúdo dos arquivos em busca de dados confidenciais ou informações críticas do sistema.

- **<u>Arquivos de configuração</u> :**

Como mencionamos anteriormente, UNIX/LINUX registra informações de configuração em arquivos específicos. Esses arquivos têm grande importância no gerenciamento de recursos do sistema. Eles devem ser editados com cautela e as permissões de acesso devem ser configuradas em bases tão restritivas quanto possível. Além disso, é recomendável que eles sejam objeto de procedimentos regulares de verificação e auditoria.

Ainda que possam ocorrer algumas diferenças de nomenclatura e localização conforme a versão de UNIX ou distribuição LINUX considerada, enumeramos a seguir alguns dos arquivos de configuração mais importantes :

- **rhosts e hosts.equiv - Localizados em root (/) e em /etc respectivamente, estes arquivos relacionam os hosts com os quais a máquina local tem um relacionamento de confiança estabelecido.**

- **inittab – Localizado em /etc, este arquivo contém regras que orientam a inicialização do sistema. Ele é interpretado pelo programa init que carrega programas, executa scripts e define níveis de execução do sistema.**

- **inetd.conf e services – Localizados em /etc, estes arquivos contém informações sobre a configuração de interfaces de rede, vinculações de protocolos e inicialização dos serviços, conhecidos como daemons em ambiente UNIX/LINUX.**

- **resolv.conf – Localizado em /etc, este arquivo informa o servidor de nomes que deve ser utilizado para resolução de endereços.**

- **hosts – Localizado em /etc, este arquivo é usado para a resolução de nomes de hosts em endereços IP correspondentes. A entradas em hosts podem conter, além do endereço IP e do nome do host, alias opcionais para**

facilitar sua localização na rede local ou simplificar o acesso via web. Uma entrada típica no arquivo hosts seria

193.107.65.11 server1.apollo.com.br server1 www.apollo.com.br

- **hosts.deny – Localizado em /etc, este arquivo contém entradas compostas por pares de hosts e serviços. Quando listado aqui, o serviço não permitirá a conexão do host especificado na mesma entrada. Uma entrada típica em hosts.deny seria**

ftp : host_A

significando que o daemon ftp não deve aceitar conexões de host_A.

- **hosts.allow – Localizado em /etc, este arquivo complementa as configurações registradas no arquivo hosts.deny, mantendo entradas similares que definem pares de daemons e hosts clientes que podem ser conectados. A inclusão de uma entrada neste arquivo apenas permite a conexão, mas não a estabelece. Permanecem válidas outras restrições de acesso associadas a cada serviço específico, como autenticação em serviços ftp. Tanto hosts.deny quanto hosts.allow aceitam os parâmetros ALL e EXCEPT para, respectivamente, designar qualquer host e excluir um determinado host da restrição configurada.**

- **exports – Localizado em /etc, este arquivo contém entradas que definem as configurações de compartilhamento de diretórios na rede, indicando permissões e níveis de acesso. Uma entrada típica em exports seria**

/home/planilhas contab*.finanças (rw)

significando que todo host cujo nome comece com a string contab e pertença ao domínio finanças tem acesso de leitura e gravação (rw) ao diretório /home/planilhas.

- **Serviços de acesso remoto :**

Muitas características do UNIX derivaram naturalmente do seu emprego como plataforma para servidores em ambientes onde as necessidades de compartilhamento de informações são acentuadas, como universidades ou grandes empresas. Uma dessas características foi a incorporação pelo sistema de uma série de serviços de acesso remoto. Tais serviços têm como um de seus principais objetivos facilitar a conexão de usuários para consulta ou recuperação de arquivos. São exemplos o telnetd, o ftpd e os chamados serviços " r ", rlogin, rsh e rexec.

Os serviços " r " em especial são responsáveis por falhas de segurança diversas. A dispensa de uma autenticação completa em determinadas circunstâncias é uma dessas falhas. A possibilidade de conceder acesso com direitos de root a um invasor mediante a simples inserção de comandos é outra ainda mais grave, embora de ocorrência menos freqüente.

Na extensa lista de vulnerabilidades associadas a esses serviços notamos que muitas delas afetam as diversas versões ou distribuições de UNIX/LINUX em graus diferenciados. Desabilitar os serviços vulneráveis é uma recomendação sensata, mas que nem sempre pode ser cumprida. Uma alternativa para prevenir ou controlar os efeitos dessas vulnerabilidades é uso de SSH – Secure Shell, um protocolo de autenticação para sessões remotas que emprega criptografia forte.

2. **Microsoft Windows NT:**

Microsoft Windows NT é o sistema operacional desenvolvido e comercializado pela Microsoft para uso corporativo. Sua primeira versão comercial, a versão 3.1, foi distribuída a partir de 1993. Em 2000, a versão que substituiu o Windows NT 4.0 foi trazida ao público com a denominação de Windows 2000. Windows 2000 incorporou diversos aperfeiçoamentos em relação ao Windows NT 4.0, mas devido às semelhanças acentuadas entre as duas versões do Windows NT, nos referimos a

ambas como Windows NT, indicando a versão 4.0 ou Windows 2000 apenas quando for necessário apontar características específicas de uma delas.

Beneficiando-se de características como interfaces gráficas bem elaboradas e uma interação simples e eficiente com aplicações desktop, e apoiado ainda pelo sucesso comercial da Microsoft, o Windows NT conta atualmente com uma ampla base instalada, tendo presença marcante em ambientes corporativos e competindo com o UNIX pela hegemonia como plataforma para aplicações de rede. Neste campo, a Microsoft oferece produtos que se integram de maneira muito eficiente com o Windows NT, como o servidor web Internet Information Server e o servidor de correio Microsoft Exchange.

Comercialmente, o Windows NT é apresentado como dois produtos distintos : o Windows NT Server, usado em servidores e o Windows NT Workstation, para uso em estações de trabalho. Com o lançamento do Windows 2000, esses produtos recebram as denominações Windows 2000 Server e Windows 2000 Professional, respectivamente. Os dois produtos têm em comum algumas de suas principais características, dentre elas o compartilhamento de recursos, a portabilidade para várias plataformas de hardware e a execução de processos em ambiente multitarefa. Embora o Windows NT Server possua funcionalidades mais amplas e venha a diferenciar-se do Workstation por traços como o suporte a múltiplas conexões simultâneas e a possibilidade de ser usado como plataforma para aplicações e serviços, eles são em essência o mesmo sistema operacional.

Dentre as características do Windows NT, as seguintes interessam diretamente a este trabalho :

- **Estrutura de Domínios :**

  As redes compostas por hosts Windows NT podem ser organizadas em dois tipos de estrutura lógica, os workgroups, ou grupos de trabalho, e os domínios.

Os domínios são estruturas hierárquicas, nas quais o acesso aos recursos compartilhados é controlado de forma centralizada por um servidor que detém o status de PDC - Primary Domain Controller, apoiado por um número variável de servidores designados como BDCs - Backup Domain Controllers. Na verdade, o emprego de BDCs não é obrigatório. O único requisito indispensável para a criação de um domínio é a existência de um PDC. O PDC do domínio mantém um banco de dados de informações de segurança denominado SAM – Security Accounts Manager. O SAM contém informações e regras referentes às contas de usuários e de grupos. Esse banco de dados é periodicamente replicado a partir do PDC para os BDCs do mesmo domínio.

O processo de login, que é denominado logon na documentação da Microsoft, é conduzido com base nas informações que constam do SAM. Quando um usuário faz logon num domínio, sua solicitação é recebida pelo PDC ou por um dos BDCs disponíveis, que deverá autenticá-lo. A possibilidade de autenticar um usuário confere aos BDCs uma utilidade adicional, além de simplesmente oferecer redundância. De fato, a carga representada pelo processo de logon dos usuários pode ser balanceada pela distribuição conveniente de vários BDCs na rede.

Pela natureza crítica das informações que contém, o SAM é o alvo de diversos ataques conhecidos. Embora seja vantajosa sobre muitos aspectos, a replicação do SAM para os BDCs pode contribuir para aumentar a exposição desse componente.

Além do controle centralizado, que contribui para um modelo de segurança mais consistente, a estrutura de domínios oferece outras vantagens, tais como maior escalabilidade, possibilidade de refletir estruturas organizacionais ou limites geográficos e flexibilidade administrativa.

Os workgroups, por sua vez, são formados por hosts reunidos em uma estrutura mais simples, não hierárquica, na qual cada host controla isoladamente o acesso aos seus próprios recursos. Sua estrutura reflete o modelo de rede ponto-a-ponto, onde cada nó pode atuar a um só tempo como cliente e servidor. Num

workgroup, cada host tem seu próprio banco de dados de segurança, autenticando localmente os usuários. Os workgroups exigem um esforço administrativo consideravelmente menor que os domínios e podem ser uma escolha adequada para redes de pequeno porte. Contudo, o controle descentralizado eventualmente implica em uma segurança mais frágil.

- **Relacionamentos de confiança :**

Diversos domínios podem ser vinculados entre si por meio de relações de confiança. As relações de confiança são configuradas entre domínios diferentes para permitir que usuários autenticados por um domínio confiável possam acessar recursos em um domínio confiante sem a necessidade de uma conta específica ou de um novo login.

Cada relacionamento de confiança é configurado como uma relação unidirecional, não recíproca e não transitiva. Isto significa que o fato de um domínio A estar configurado para confiar em um domínio B não implica que o domínio B irá confiar no domínio A. Tal condição exigiria a configuração de um segundo relacionamento de confiança, no qual os papéis fossem invertidos, isto é, o domínio B confiasse explicitamente no domínio A. Da mesma forma, o relacionamento de confiança estabelecido entre dois domínios não é estendido a outros domínios pelo fato de haverem relacionamentos de confiança adicionais. Assim, se o domínio A confia em B e B confia em um terceiro domínio C, não será estabelecido com isso que o domínio A confia em C. Isso só aconteceria se fosse criado um relacionamento de confiança entre os domínios A e C.

O Active Directory, recurso incorporado ao Windows 2000, altera esse quadro, estabelecendo novas características para os relacionamentos de confiança, como veremos adiante.

Os relacionamentos de confiança simplificam o gerenciamento de contas e direitos de acesso em redes complexas, proporcionando uma estrutura mais escalável e possibilitando o emprego de métodos de administração distribuída.

Contudo, eles representam uma vulnerabilidade potencial, uma vez que uma exploração bem conduzida dessa característica pode deixar um atacante em condições de acessar recursos num domínio confiante.

- **Sistema de arquivos NTFS :**

Embora reconheça e possa operar em unidades e partições formatadas por sistemas de arquivos tradicionais como FAT e HPFS, Windows NT dispõe de um sistema de arquivos próprio, denominado NTFS – NT File System. O NTFS foi desenvolvido especialmente para uso com este sistema operacional. Ele é um sistema de arquivos sofisticado e seguro, que incorpora mecanismos avançados de controle e recuperação de erros.

No NTFS, a estrutura da MFT – Master File Table, que corresponde à FAT – File Allocation Table de outros sistemas de arquivos, foi definida de modo a orientar com facilidade o acesso a arquivos e diretórios. O sistema mantém uma cópia de segurança da MFT em uma área do disco fisicamente afastada da localização da MFT original. Isso diminui os riscos de perda da MFT. As operações de gravação no NTFS também são mais seguras em razão de um mecanismo de controle baseado em logs de transação. Esse mecanismo faz com que cada operação seja registrada antes de sua execução, permanecendo o registro em aberto até que a operação seja concluída. Caso haja uma falha inesperada, decorrente de corte da energia elétrica ou outra causa fortuita, e a transação seja interrompida, ela poderá ser retomada com base nas informações do log, garantindo assim a integridade dos dados.

Naturalmente, a complexidade dessas funções implica num overhead relativamente alto. A Microsoft recomenda uma análise da viabilidade do emprego do NTFS baseada em características de hardware para sistemas cujos recursos sejam limitados, a fim de preservar uma performance adequada.

Contudo, no que diz respeito à segurança de dados, a característica mais importante do NTFS é a possibilidade de estabelecer restrições de acesso a nível

de arquivo. Essas restrições são definidas pelas chamadas permissões NTFS, cujo emprego é detalhado a seguir. Na prática, as permissões NTFS atribuídas sobre um arquivo ou diretório constituem-se em atributos especiais ou propriedades dos arquivos.

- **Restrições de acesso :**

As restrições de acesso no Windows NT são implementadas em dois níveis : compartilhamentos e sistema de arquivos.

No primeiro nível, Windows NT permite a atribuição de permissões de compartilhamento, que definem o grau de acesso que um determinado usuário ou grupo terá sobre um recurso compartilhado quando acessá-lo através da rede. As permissões de compartilhamento aplicam-se a recursos compartilhados, tais como diretórios e impressoras de rede. Elas não podem ser aplicadas a arquivos individuais. Elas especificam que operações o usuário ou grupo poderá executar com o recurso em questão. Essas operações variam conforme o recurso considerado. Assim, para arquivos e diretórios existem as seguintes permissões :

- **No access (Sem acesso)**
- **Read (Leitura)**
- **Write (Gravação)**
- **Full Control (Controle Total)**

Já para impressoras compartilhadas as permissões são as seguintes :

- **No access (Sem acesso)**
- **Print (Imprimir)**
- **Manage documents (Gerenciar documentos)**
- **Full Control (Controle Total)**

As restrições impostas pelas permissões de compartilhamento limitam exclusivamente o acesso através da rede, ou seja, não fornecem controle efetivo quando o recurso é acessado localmente.

As restrições de sistema de arquivos são estabelecidas por meio de permissões NTFS, assim chamadas por serem suportadas por esse sistema de arquivos, sendo sua implementação possível apenas em partições formatadas com NTFS. As permissões NTFS oferecem controle sobre arquivo individuais e diretórios, inclusive para acesso local a esses recursos.

O conjunto de permissões NTFS é mais elaborado que o das permissões de compartilhamento, combinando um certo número de permissões padrão em permissões aplicáveis a diretórios, permissões individuais para arquivos e permissões de acesso especial.

As permissões NTFS padrão são as seguintes :

- **Read (R) - Leitura**
- **Execute (X) - Execução**
- **Write (W) - Escrita ou gravação**
- **Delete (D) - Exclusão**
- **Change permissions (P) – Alterar permissões**
- **Take ownership (O) – Tomar posse**

As tabelas apresentadas a seguir demonstram como as permissões NTFS padrão são combinadas em permissões mais complexas, restringindo a execução de operações sobre diretórios e arquivos :

Permissões para pastas ou diretórios		
Tipo	**Combinação**	**Ações permitidas**
Sem acesso	**-**	**Nenhuma.**
Listar	**R**	**Examinar o diretório, seus subdiretórios e ter acesso aos nomes dos arquivos neles contidos.**

Permissões para pastas ou diretórios (cont.)		
Leitura	**R + X**	**Ler o conteúdo de arquivos e executar programas.**
Adição	**W + X**	**Inserir novos arquivos no diretório, sem ler seu conteúdo ou alterá-los.**
Adição e leitura	**R + W + X**	**Inserir novos arquivos no diretório, ler seu conteúdo e alterá-los.**
Alteração	**R + W + X + D**	**Inserir novos arquivos no diretório, ler seu conteúdo, alterá-los e excluí-los.**
Controle total	**R + W + X + D + O**	**Inserir novos arquivos no diretório, ler seu conteúdo, alterá-los, excluí-los, tomar posse deles e alterar as permissões de acesso estabelecidas.**

Permissões para arquivos		
Tipo	**Combinação**	**Ações permitidas**
Sem acesso	**-**	**Nenhuma.**
Leitura	**R + X**	**Ler o conteúdo do arquivo e, se for um programa, executá-lo.**
Alteração	**R + W + X + D**	**Ler o conteúdo do arquivo, alterá-lo e excluí-lo.**
Controle total	**R + W + X + D + O**	**Ler o conteúdo do arquivo, alterá-lo, excluí-lo, tomar posse dele e alterar as permissões de acesso estabelecidas.**

As permissões especiais sobre arquivos ou pastas são obtidas a partir da associação das permissões padrão em combinações diferentes das mencionadas nas tabelas acima. Elas fornecem graus de controle adicionais sobre o acesso a arquivos e pastas.

As permissões NTFS podem ser combinadas ainda com as permissões de compartilhamento, prevalecendo em cada caso a combinação de direitos mais restritiva. Esse amplo conjunto de possibilidades faz com que o Windows NT seja capaz de oferecer grande precisão e flexibilidade no controle de acesso sobre pastas, arquivos e outros recursos. Entretanto, a relativa complexidade de configuração de direitos, decorrente sobretudo do grande número de opções e dos métodos usados para combiná-las em permissões efetivas, exige uma administração metódica e cuidadosa. A situação exposta a seguir exemplifica essa necessidade.

Quando um recurso é compartilhado, Windows NT por padrão atribui sobre ele a permissão Full Control para o grupo Everyone. Como este grupo inclui todos os usuários, isto significa que na prática são concedidos direitos irrestritos a qualquer usuário sobre o compartilhamento recém-criado. Supostamente, essa característica existe para prevenir que um recurso torne-se inacessível, caso não haja nenhum usuário ou grupo com direitos suficientes para controlar o acesso a ele. Contudo, isso sempre obriga o administrador a reconfigurar o acesso definido automaticamente pelo sistema operacional, removendo os direitos concedidos ao grupo Everyone e atribuindo outros conforme as exigências da situação. Fica estabelecida com isso uma condição insegura, desde que uma falha nesse procedimento poderá deixar um compartilhamento aberto ao acesso de qualquer usuário.

Como vimos, o acesso a recursos da rede pode ser controlado pela configuração de restrições de acesso baseada em permissões de compartilhamento e permissões NTFS. Além dessa possibilidade, Windows NT fornece meios que podem ser usados para restringir o acesso do usuário a informações críticas e até mesmo ao uso de programas e interfaces de configuração em suas estações de trabalho. As diretivas de contas, as diretivas de sistemas ou policies e os perfis de usuário reúnem uma grande variedade de opções por meio das quais é possível estabelecer uma política eficaz de manutenção e renovação de senhas, padronizar interfaces, definindo um ambiente de trabalho consistente com as necessidades de segurança, e controlar as ações dos usuários a nível de desktop. A profusão de

detalhes envolvidos no seu uso leva-me a optar por não descrever aqui esses métodos. Cumpre assinalar, entretanto, que eles devem ser considerados como um recurso efetivo para a manutenção da segurança.

- **Suporte a múltiplos protocolos :**

O Windows NT oferece suporte nativo a uma grande variedade de protocolos, dentre eles o TCP/IP, o NetBEUI e IPX/SPX, este último distribuído no Windows NT com o nome de NWLink. Esta característica empresta grande versatilidade ao sistema operacional, possibilitando seu uso em ambientes heterogêneos e sua integração com sistemas diversos.

Cabe mencionar aqui que a utilização de múltiplos protocolos em sistemas Windows NT deve seguir um planejamento cuidadoso, em que se considere não só as necessidades de comunicação, mas também as implicações sobre a segurança que isso poderá trazer. Uma vez mais é válida a recomendação no sentido de que não sejam habilitados protocolos ou serviços que não serão utilizados, a fim de restringir as possibilidades de exploração.

As funções de roteamento que podem ser implementadas em servidores NT com a habilitação do RIP devem ser também cuidadosamente analisadas sob a perspectiva dos efeitos na segurança do host e das redes conectadas a ele. O protocolo RIP serve de base para técnicas de ataque ao roteamento mencionadas no capítulo 3.

Além disso, os protocolos suportados pelo Windows NT para uso em conexões de longa distância, como o PPP e o SLIP apresentam características de segurança próprias. É importante conhecê-las a fim de avaliar os efeitos de sua utilização.

O SLIP – Serial Line Protocol - é um protocolo de WAN utilizado pelo Windows NT apenas para conexão como cliente de hosts UNIX. Um servidor NT não aceita conexões baseadas em SLIP. Essa limitação reduz o risco decorrente das deficiências de segurança desse protocolo, que não suporta criptografia na

autenticação de usuários. Entretanto, seu emprego num ambiente de rede heterogêneo pode ainda representar uma vulnerabilidade importante, na medida em que expõe senhas em texto claro na rede .

Já o PPP – Point-to-Point Protocol – constitui-se no protocolo padrão do Windows NT para suporte a conexões remotas. Ele traz diversas vantagens sobre o SLIP, das quais a mais relevante no que toca à segurança da rede é o suporte à autenticação por meio de senhas criptografadas.

- **NetBIOS e WINS :**

O NetBIOS foi criado em 1983 no âmbito de um projeto da IBM para desenvolver um protocolo capaz de viabilizar a comunicação entre aplicações executadas em diversos hosts conectados. O NetBIOS é um protocolo não roteável, simples e rápido, com overhead muito baixo. Por essa razão, é utilizado pelo Windows NT para suporte a funções básicas, como acesso remoto a arquivos, transferência de dados e impressão em rede.

O NetBIOS engloba uma API – Application Programming Interface - e um protocolo de controle de sessão. A interface de programação disponibiliza um conjunto de funções e comandos, permitindo que aplicações utilizem uma conexão existente para transmitir e receber dados. As funções do NetBIOS são dessa forma implementadas sobre conexões gerenciadas pelos protocolos das camadas inferiores. Quando o protocolo da camada de transporte que fornece ao NetBIOS os serviços de conexão é o TCP/IP, emprega-se o termo NetBIOS sobre TCP/IP.

NetBIOS define um sistema de nomeação plano ou não hierárquico, no qual cada host recebe um nome único com extensão de até dezesseis bytes. Os nomes NetBIOS podem ser usados para a localização de recursos em uma rede de proporções limitadas. Numa rede mais ampla ou num sistema de redes interconectadas, as possibilidades do NetBIOS são consideravelmente reduzidas. Como o sistema de nomeação NetBIOS não é hierárquico, todos os hosts detém status equivalente. Embora seja possível implementar um nível elementar de

organização de espaço de nomes, reunindo hosts em grupos denominados escopos NetBIOS, não é possível o estabelecimento de estruturas organizadas em níveis que facilitem a localização e o gerenciamento.

NetBIOS utiliza broadcasts para registrar nomes de host e para localizar computadores na rede. Em condições usuais, os broadcasts têm seu alcance limitado ao segmento de rede local, não ultrapassando a fronteira representada pelos roteadores que ligam a rede local a outras redes. Quando um computador necessita localizar e estabelecer conexão com outro host baseado em seu nome NetBIOS, ele envia um broadcast de consulta que é respondido pelo host de destino com seu endereço IP.

A resolução de nomes NetBIOS, ou seja, sua conversão em endereços IP pode ser feita também por meio de arquivos de texto contendo entradas que vinculem em pares os nomes NetBIOS aos endereços IP correspondentes. Esses arquivos, disponíveis em cada host, devem obedecer aos padrões de denominação e localização estabelecidos pelo sistema operacional. No caso do Windows NT, o arquivo de resolução de nomes NetBIOS é denominado LMHOST e localiza-se em WINNT\SYSTEM32\DRIVERS\ETC.

O emprego de arquivos de nomes possibilita que hosts posicionados em redes remotas sejam localizados por seus nomes NetBIOS. Nesses casos, o host de origem não dependerá de broadcasts limitados à rede local para obter o endereço IP do host de destino. Ele examinará o conteúdo do arquivo LMHOSTS. Havendo ali uma entrada que associe o nome NetBIOS do host de destino a um endereço IP, o host de origem terá a informação necessária para solicitar a abertura de uma conexão.

Um benefício adicional do emprego de arquivos de texto como LMHOSTS para a resolução de nomes NetBIOS é a redução de tráfego na rede, uma vez que substitui o método de resolução baseado em broadcasts. Existem, porém, alguns problemas relacionados ao seu uso. A mais importante delas decorre do fato de que os arquivos LMHOSTS são estáticos, havendo a necessidade de atualização manual a

cada alteração ocorrida no ambiente, como a inclusão de novos hosts, por exemplo. Em uma rede com grande número de hosts, isso pode representar um encargo administrativo difícil de gerenciar.

A utilização de um servidor de nomes é outra forma de estender o alcance do NetBIOS, possibilitando sua utilização em ambientes com mais de um segmento de rede. Havendo a disponibilidade de um NBNS – NetBIOS Name Server , o host de origem enviará sua consulta em um pacote unicast direcionado ao servidor de nomes. Caso disponha da informação necessária à resolução solicitada, o servidor de nomes responderá enviando também um pacote unicast ao host de origem. Sendo assim, o emprego de servidor de nomes NetBIOS também contribui para limitar a sobrecarga de tráfego na rede, reduzindo a ocorrência de consultas em broadcasts.

Windows NT utiliza uma implementação de servidor de nomes NetBIOS denominada WINS – Windows Internet Name Service. WINS é instalado em servidores NT como um serviço adicional. Os clientes da rede devem ser configurados então para utilizarem WINS como um método de resolução de nomes NetBIOS. Como parte dessa configuração, é especificado também o servidor WINS que cada cliente deverá consultar. Ao inicializar, o cliente registra seu nome NetBIOS junto ao servidor WINS configurado. A base de dados do WINS é assim atualizada dinamicamente, restringindo a possibilidade de erros e simplificando a administração.

A principal vulnerabilidade associada ao emprego do WINS como método de resolução de nomes decorre do fato de que o registro do nome NetBIOS de um host na base de dados do servidor é feito sem qualquer tipo de autenticação ou verificação, favorecendo dessa forma a ocorrência de ataques de personificação como IP spoofing, que é apresentado no Capítulo 3 deste trabalho. Num ataque de IP spoofing, um host configurado com um nome NetBIOS falso poderá registrar esse nome junto ao WINS, disponibilizando essa informação para os demais hosts da rede. WINS possui métodos para tratar nomes duplicados, mas eles limitam-se a evitar que um host registre um nome já utilizado por outro host. Naturalmente, o

atacante interessado em simular a identidade de outro host pode até mesmo evitar essa situação aproveitando-se de uma interrupção normal de funcionamento do host personificado ou mantendo o host fora do ar mediante o uso de uma técnica de negação de serviço. A utilização de entradas fixas, inseridas manualmente na base de dados WINS representa uma forma de exercer algum controle sobre essa vulnerabilidade. Ela pode ser aplicada para hosts que executem funções críticas na rede, mas dificilmente um administrador cogitaria em utilizar entradas fixas para todos os hosts da rede, já que estaria abrindo mão assim da atualização dinâmica da base de dados, uma das principais funcionalidades do WINS.

- **Serviços de acesso remoto :**

O Windows NT fornece um serviço que permite a conexão de clientes remotos a um servidor Windows NT Server através de links de longa distância, como linhas discadas e conexões ISDN ou X.25. O RAS – Remote Access Services – é instalado como um serviço em um Windows NT Server. Após a instalação, o serviço deve ser vinculado a um ou mais protocolos. O RAS suporta os protocolos TCP/IP, NWLink e NetBEUI.

Quando conectados a um servidor RAS, os clientes remotos podem ter acesso a recursos não somente nesse servidor, mas em outros hosts da rede local. Por atuar na ligação da rede local com nós remotos, o RAS impõe requisitos de segurança compatíveis com os riscos que sua utilização importa. O serviço conta com recursos próprios de segurança e controle, que descrevemos a seguir.

- **O processo de autenticação de usuários que se conectam à rede por meio do RAS permite o emprego opcional de senhas criptografadas. Três configurações estão disponíveis e devem ser selecionadas conforme a situação:**

1. **Aceitar qualquer autenticação inclusive texto claro – Esta configuração determina que sejam utilizadas senhas não criptografadas. Ela é recomendável apenas quando a segurança de logon não é um requisito**

ou quando os hosts envolvidos não suportam qualquer método de autenticação criptografada.

2. **Aceitar somente autenticação criptografada – Esta configuração utiliza padrões não proprietários para a criptografia das senhas. Deve ser usada quando há necessidade de preservar a confidencialidade das senhas, mas um dos hosts envolvidos é um servidor ou cliente não-Microsoft.**

3. **Aceitar somente autenticação criptografada Microsoft – Esta configuração emprega o MS-CHAP – Microsoft Challenge Handshake Authentication Protocol – e exige que ambos os hosts envolvidos sejam servidores ou clientes Microsoft.**

- **As restrições e diretivas de segurança implementadas no domínio para contas de usuários e de grupos permanecem funcionais quando um usuário faz logon no domínio por meio de um servidor RAS. Como o serviço RAS executa uma autenticação do usuário remoto baseada no SAM do servidor NT e em permissões de discagem configuradas para aquele usuário, o conjunto de restrições e privilégios associado à conta do usuário mantém sua consistência. Esse modelo de logon em rede única, como referenciado na documentação da Microsoft, simplifica a administração e incrementa a segurança do processo.**

- **RAS oferece a possibilidade de configurar verificações de segurança aplicáveis às chamadas recebidas. Quando configuradas, essas verificações funcionam interrompendo a conexão logo após o recebimento da chamada, discando de volta para o número que executou a chamada ou para outro número predefinido. Esse mecanismo de callback, como é chamado, assegura que a conexão está sendo feita de um local conhecido ou confiável.**

- **RAS permite a utilização de um host de segurança intermediário, posicionado entre o cliente remoto e o servidor RAS. Esse host irá executar uma autenticação adicional do cliente antes que a conexão com o servidor RAS**

possa ser feita. A utilização de um host de segurança intermediário não elimina a necessidade de autenticação imposta pelo servidor RAS, mas acrescenta um nível de segurança extra à conexão.

- **As autenticações via RAS geram informações passíveis de serem registradas em logs de auditoria. O mecanismo de auditoria do Windows NT é descrito em detalhes mais adiante neste capítulo.**

- **Um servidor RAS dotado de uma conexão contínua com a Internet pode ser configurado para utilizar o PPTP – Point-to-Point Tunelling Protocol – encapsulando os pacotes TCP/IP, IPX ou NetBEUI enviados para a grande rede em pacotes desse protocolo. O PPTP criptografa os dados transmitidos, assegurando sua integridade e confidencialidade. Os pacotes recebidos são filtrados no adaptador de rede conectado à Internet. O adaptador aceitará apenas os pacotes PPTP destinados à rede local ou interna. Pacotes de outros protocolos serão descartados. Essa técnica é a base das VPNs – Virtual Private Networks-, uma solução bastante difundida e eficaz, capaz de fornecer um intercâmbio seguro de dados pela Internet, além da redução dos custos de comunicação. As VPNs são descritas no capítulo 4**.

- **Auditoria :**

Windows NT dispõe de uma ferramenta de auditoria integrada ao sistema operacional, por meio da qual é possível observar a ocorrência de determinados eventos, registrando informações em um arquivo de log para análise posterior.

A auditoria do Windows NT é definida para um domínio como uma diretiva que especifica quais os eventos que deverão ser auditados e como eles serão registrados. Os eventos que podem ser auditados pelo serviço de auditoria do Windows NT são os seguintes :

  - **Acesso e alterações em diretórios e arquivos**
  - **Acesso e alterações em impressoras**

- **Logon e logoff de usuários**
- **Desligamentos e reinicializações do sistema**
- **Alterações em contas de usuários e grupos**
- **Alterações em diretivas de segurança**

A auditoria aplicada sobre arquivos e diretórios requer que os mesmos estejam em uma partição NTFS.

A auditoria irá registrar cada evento em um arquivo denominado log de segurança, informando a ação executada, o êxito ou falha da ação, o usuário que a executou ou tentou executar e a data e hora em que isso ocorreu. O exame do log se segurança pode ser facilitado pela utilização de mecanismos de filtragem que separam eventos específicos em meio a um grande número de outros eventos. Os filtros disponíveis permitem separar eventos por intervalo de tempo, por tipo, por usuário e por várias outras categorias.

O log de segurança é armazenado localmente, o que representa uma limitação importante, especialmente quando comparado com o syslog do UNIX. Caso o host venha a sofrer um comprometimento de segurança e um atacante obtenha acesso ao sistema com direitos apropriados, o log poderá ser alterado de modo a encobrir a invasão.

Como dissemos, cada evento pode ser auditado em bases de êxito ou falha, isto é, o administrador pode optar por registrar no log de segurança as ações executadas com sucesso ou aquelas que consistem em tentativas de acesso a recursos bloqueadas pelas restrições vigentes. Este segundo tipo de registro é de grande auxílio na observação das tentativas de acesso não autorizado.

Além de permitir o acompanhamento de ocorrências relacionadas à segurança do sistema, o que geralmente é feito auditando-se eventos que resultam em falhas, a auditoria do Windows NT é útil também para monitorar tendências de utilização de recursos. Nesse caso, devem ser observadas as ações que têm êxito e analisadas sua origem e freqüência, registradas ao longo de um período representativo.

Deve-se sempre ter em mente que, a par de sua grande utilidade, a auditoria envolve custos, representados pela sobrecarga imposta ao sistema para registrar continuamente os eventos relacionados na diretiva da auditoria, além do esperado crescimento do arquivo de log. Esse crescimento pode ser exagerado quando é registrado o êxito de eventos, normalmente muito mais freqüente do que as falhas, ou quando a auditoria é realizada sobre um número muito grande de eventos. Essas são algumas das razões pelas quais a configuração de auditorias não é permitida a usuários comuns, constituindo-se num privilégio concedido aos administradores do sistema.

Ainda assim, a auditoria representa um recurso extremamente valioso no planejamento e condução de uma política de segurança. Por meio dela o administrador de rede pode obter informações detalhadas sobre a utilização de recursos e detectar tentativas de acesso não autorizado. A utilidade dessas informações é muito ampla, podendo servir tanto para subsidiar ações antecipadas capazes de evitar futuras violações da segurança quanto para definir responsabilidades administrativas sobre uso indevido de recursos e outras ocorrências dessa natureza.

- **Necessidade de atualizações :**

A descoberta de vulnerabilidades e a distribuição periódica de correções e atualizações constituem-se em eventos normais no ciclo de utilização de qualquer sistema operacional de rede. No caso do Windows NT esta necessidade tem sido particularmente marcante. Desde a primeira versão foram observadas diversas falhas de segurança, algumas de considerável gravidade. Esses problemas mereceram correções por parte da Microsoft, a maioria delas oportunas e eficazes. A Microsoft tem distribuído correções em pacotes denominados Service Packs. É surpreendentemente comum, entretanto, a ocorrência de comprometimentos como resultado da aplicação inadequada, ou mesmo da não aplicação, dessas correções. Todo administrador que tenha sob sua responsabilidade sistemas baseados em Windows NT deve dedicar especial

atenção ao cumprimento da regra geral que recomenda manter servidores e estações de trabalho sempre atualizados com as últimas correções disponíveis.

- **Active Directory :**

O Active Directory é um serviço de diretório oferecido como parte do Windows 2000. Ele prevê interoperabilidade com outros sistemas operacionais e suporte a funções avançadas como distribuição automatizada de software e suporte a certificados digitais.

As interações entre o Active Directory e diretórios em outras redes são gerenciadas pelo LDAP – Lightweight Directory Access Protocol, um conjunto de protocolos de padrão aberto projetado para dar suporte à troca de informações entre bancos de dados de diretório em sistemas operacionais diferentes ou não. O Active Directory prevê a possibilidade de apoiar as interações entre diretórios com o emprego de conexões seguras baseadas em SSL - Secure Sockets Layer. O uso de mecanismos especiais de autenticação, como uma implementação do Kerberos, pode também ser configurado para fornecer níveis de segurança adicionais, uma exigência frequente em operações de e-commerce, por exemplo.

O Active Directory representa uma iniciativa da Microsoft para estabelecer um novo modelo de gerenciamento integrado de recursos, distinto e aperfeiçoado em relação àquele oferecido pelo Windows NT 4.0, que dependia de relações de confiança para a integração de domínios. Com o Active Directory os relacionamentos de confiança são substituídos por uma representação unificada do conjunto de recursos disponíveis em vários domínios. Embora beneficie a administração, o Active Directory apresenta vulnerabilidades relevantes. Em ambientes onde convivem servidores Windows NT 4.0 e Windows 2000 o diretório é consultado pelos servidores NT 4.0 com RAS instalado a fim de determinar as restrições vigentes para o acesso remoto. A instalação do Active Directory prevê essa possibilidade e configura o serviço para aceitar consultas remotas. Essa possibilidade pode ser explorada por um usuário remoto para a

obtenção de informações críticas do sistema. Pode ser usado para isso um utilitário denominado ldp, capaz de examinar o conteúdo do diretório.

3. **<u>Novell Netware</u>:**

NetWare é um sistema operacional de rede desenvolvido e comercializado pela Novell Inc. a partir de 1981. Sua origem remonta ao início da década de oitenta. Nessa época, as mídias de armazenamento fixas eram caras e de capacidade limitada. Os custos envolvidos no armazenamento de dados eram consideravelmente altos e, dependendo do porte da organização, a necessidade de dotar cada estação de trabalho de um disco rígido próprio podia representar gastos de milhares de dólares. O compartilhamento de um único disco rígido por vários computadores interconectados surgiu como uma alternativa atraente para fornecer capacidade de armazenamento com custos reduzidos. A Novell Data Systems, que se dedicava então à produção de computadores em Utah, nos Estados Unidos, deu início ao desenvolvimento de um sistema que permitisse que um disco rígido fosse utilizado por mais de um usuário ao mesmo tempo. O sistema, criado por uma equipe que contava com a participação de Drew Major e Dale Neibaur entre outros, obteve excelentes resultados, alcançando extraordinário sucesso comercial e incorporando subseqüentemente novos aperfeiçoamentos e funcionalidades. A versão 2.0 do sistema, lançada em 1985 trazia otimizações que lhe permitiam alto rendimento com processadores 80286, os mais avançados da época. O NetWare rapidamente impôs-se como o sistema operacional de rede mais utilizado na área empresarial e deu à Novell a hegemonia do mercado, com uma antecedência de quase dez anos em relação às suas concorrentes mais significativas. A versão 4.11 recebeu o nome comercial de intraNetWare.

Uma série de equívocos comerciais que se tornaram históricos e a pressão de concorrentes como o Windows NT e, mais recentemente, o Linux, foram responsáveis pela perda dessa posição. Atualmente, o emprego do NetWare como sistema operacional de rede apresenta uma forte tendência decrescente. Contudo, sua análise no corpo deste trabalho justifica-se por duas circunstâncias. Em primeiro lugar porque o sistema oferece um modelo de gerenciamento extremamente

consistente e avançadas funções de segurança. Essas características, além de garantir ao NetWare a reputação de sistema confiável e seguro, vieram a orientar o aperfeiçoamento de técnicas de controle e segurança empregadas em outros sistemas operacionais. Em segundo lugar, a ampla base instalada do NetWare, sobretudo em grandes corporações e órgãos públicos, coloca-o ainda como um dos sistemas operacionais de rede mais utilizados, não obstante sua presença no mercado diminuir a cada dia.

As circunstâncias que deram origem ao NetWare condicionaram a evolução do sistema por um longo período e, em muitos aspectos, fazem sentir-se ainda hoje. Elas determinaram características que fizeram do NetWare um sistema operacional confiável e seguro. Expomos a seguir algumas dessas características :

- **Arquitetura cliente – servidor :**

  Uma rede NetWare é marcada pela definição das funções de cliente e servidor de maneira acentuada. De fato, NetWare limita as possibilidades de compartilhamento de recursos a nível de estação de trabalho e centraliza nos servidores as funções de controle de acesso.

  Essa característica restringe a flexibilidade do sistema operacional, impedindo que muitos recursos das estações de trabalho, como arquivos e diretórios, sejam compartilhados na rede. Em contrapartida, o modelo de segurança determinado pelo sistema operacional apresenta grande consistência e contribui para o estabelecimento de condições de segurança muito sólidas. A concentração de recursos e funções nos servidores permite ao administrador de uma rede NetWare focalizar esforços, com consequências favoráveis sobre o controle e a segurança. A criação de pontos únicos de falha é satisfatoriamente compensada pela estabilidade do sistema e por mecanismos de redundância que serão mencionados adiante.

- **Novell Directory Services :**

Até a versão 3.12, o NetWare gerenciava recursos a partir de uma base de dados denominada bindery, instalada com o sistema operacional em cada servidor. O bindery controlava apenas os recursos disponíveis no servidor NetWare em que estava instalado. Num cenário de múltiplos servidores, era necessário que os usuários mantivessem uma conta em cada servidor para ter acesso aos recursos de todos eles. Evidentemente esse era um método pouco vantajoso, complicando a administração da rede e limitando sua escalabilidade. Com a versão 4.1 do sistema foi introduzido o NDS, um serviço que compunha uma base de dados, denominada Diretório, muito mais ampla que o bindery, capaz de conjugar informações sobre vários servidores em uma estrutura lógica e hierárquica denominada árvore. O NDS expandiu de forma significativa a capacidade do NetWare para suportar grandes redes, centralizando as funções de gerenciamento em um serviço diretórios altamente integrado.

O NDS é compatível com o padrão X.500 para serviços de diretórios. Ele permite acesso lógico a recursos independente de sua localização física e fornece interoperabilidade com outros tipos de rede. Ele foi responsável também por ganhos importantes de eficiência e estabilidade.

O NDS representa cada recurso da rede como um objeto ao qual são associadas propriedades específicas. Uma conta de usuário, por exemplo, é representada por um objeto Usuário, com propriedades como restrições de senha e scripts de login. Uma impressora compartilhada é representada por um objeto Impressora, com propriedades como tipo de impressora e endereço de rede. O NDS autentica usuários e controla o acesso a recursos com base na identificação dos objetos que os representam e em suas propriedades.

O acesso aos recursos é controlado no NetWare por meio de designações de trustees. Um trustee é um objeto ao qual são atribuídos direitos sobre outro objeto. Assim, para acessar um diretório compartilhado num servidor NetWare, um objeto Usuário deve ser designado como trustee daquele diretório, com os

direitos apropriados. A cada solicitação de acesso, o sistema verifica a base de dados e, examinando as propriedades do objeto que representa o recurso, determina se os direitos configurados permitem o acesso. Esse método garante um controle consistente e eficaz, a partir de uma base de dados unificada que congrega todas as informações necessárias a essa função. O NDS permite ainda a reunião de usuários em grupos, também representados por um objeto no NDS, a propagação dos direitos atribuídos ao longo da estrutura da árvore conforme regras de herança e a aplicação de filtros denominados IRFs – Inherited Rights Filters, para controlar essa propagação.

O NDS mantém uma cópia principal de sua base de dados em um servidor determinado, que é normalmente o primeiro servidor instalado na árvore, embora isso possa ser alterado. Essa cópia principal é denominada réplica master do Diretório. Réplicas do diretório são criadas automaticamente no segundo e terceiro servidores instalados na árvore. Se necessário, outras réplicas podem ser criadas em servidores instalados subsequentemente. Essas novas réplicas têm características diferentes da réplica master, sendo classificadas, conforme essas características, em réplicas read-write ou réplicas read-only. Existe ainda um quarto tipo de réplica, denominada referência subordinada, com finalidades especiais que não serão abordadas aqui.

O acesso a recursos pode ser autorizado pela consulta às informações do Diretório que constam das réplicas, mas as alterações nas propriedades dos objetos devem ser requisitadas à réplica master, que as autoriza e registra. Periodicamente, o servidor que mantém a réplica master replica as alterações do Diretório para os outros servidores que contenham réplicas. A função de replicação preserva a integridade do Diretório e acelera o acesso a recursos, permitindo que réplicas do Diretório sejam distribuídas estrategicamente na rede e posicionadas próximas aos recursos ou aos usuários que os acessam.

O êxito do NDS em prover gerenciamento a redes de grande porte, aliado a suas características de interoperabilidade, levaram-no a ser portado para outras plataformas e vendido pela Novell como um produto separado.

- **Suporte limitado aos protocolos TCP/IP :**

  Até sua versão 4.12 o NetWare não oferecia suporte nativo aos protocolos TCP/IP. Os protocolos das camadas de Internet e Transporte padrão do sistema eram respectivamente o IPX – Internetwork Packet Exchange e o SPX – Sequenced Packet Exchange. Embora o IPX seja um protocolo roteável e bastante eficiente em várias funções, ele não é tão eficiente quanto o TCP/IP para dar suporte à comunicação num ambiente complexo de redes interconectadas. A integração de uma rede exclusivamente IPX com redes TCP/IP era possível pela instalação nos servidores NetWare de um serviço especial, denominado IPX Gateway, que fazia a tradução de pacotes IPX em pacotes TCP/IP e vice-versa. Esse modelo funcionava bem em redes pequenas e médias, contribuindo inclusive para que as redes que operavam exclusivamente o IPX estivessem de certa forma preservadas de vulnerabilidades específicas do TCP/IP. Contudo, à medida que as redes expandiam-se e que o tráfego inter-redes crescia, a necessidade de suporte mais amplo ao TCP/IP acabou por impor-se. A partir da versão 5.0, o NetWare passou a fornecer uma implementação própria do TCP/IP.

- **Plataforma para aplicações avançadas :**

  As versões mais recentes do NetWare, além de oferecerem amplo suporte ao TCP/IP, funcionam como plataforma para aplicações e serviços avançados. Dentre eles estão o servidor de aplicativos IBM Websphere Application Server, o software de gerenciamento remoto de desktops Z.E.N.works, a ferramenta de groupware Novell GroupWise e outros.

  Dentre as aplicações que utilizam o NetWare como plataforma, destaca-se o Novell Border Manager. O Border Manager é uma suíte de aplicativos orientada à segurança. Estreitamente integrado com o NDS, ele provê um conjunto bastante amplo de serviços que podem ser gerenciados de maneira unificada. Dentre eles temos :

- **serviços de proxy e firewall, capazes de controlar o acesso de usuários externos a recursos da rede local, bem como o acesso de usuários internos a redes remotas e à Internet.**

- **serviços de autenticação compatíveis com vários métodos de login, como senhas, tokens, certificados digitais e processos biométricos.**

- **serviços de suporte a VPN, que possibilitam a montagem de redes privadas virtuais interligando redes remotas.**

Novell Border Manager oferece ainda recursos como cache de conteúdo web e serviços de acesso para usuários remotos.

Capítulo 3

Violações da segurança

Violações da segurança

O desenvolvimento de técnicas destinadas a comprometer serviços nas redes baseadas em TCP/IP ou a fornecer acesso não autorizado a dados manipulados no âmbito dessas redes vem acompanhando o crescimento exponencial da Internet e das redes corporativas que utilizam essa tecnologia.

Julguei oportuno descrever alguns desses métodos não pela engenhosidade que eventualmente revelam, nem tampouco pela série de outras razões menos justificáveis que já estimulou a criação de inúmeros textos sobre o tema. Na verdade, as técnicas de ataque concebidas para utilização em redes TCP/IP consideram e exploram falhas de segurança, eventuais ou intrínsecas, do conjunto de protocolos mais usado na comunicação de computadores em rede. Por isso, pareceu-me que investigar seu funcionamento, suas formas de emprego e, quando possível, os conceitos que orientaram seu desenvolvimento corresponderia a analisar possibilidades e limitações dos protocolos sob uma perspectiva eminentemente prática.

Naturalmente não pretendo elaborar um inventário das técnicas disponíveis no dia de hoje, inclusive porque amanhã certamente já existirão outras. Também não é meu propósito apresentar as técnicas mais utilizadas ou as mais eficientes. Procuro descrever aqui aquelas que me pareceram mais representativas das falhas de segurança associadas aos protocolos TCP/IP, procurando focalizar as maneiras pelas quais as características dos protocolos são aproveitadas no comprometimento da segurança.

Deve ser entendido que algumas dessas técnicas não têm seu uso restrito ao ambiente de redes TCP/IP, podendo ser empregadas em outros tipos de rede. Da mesma forma, técnicas que se apoiam em protocolos e funções da camada superior são também mencionadas, por sua associação freqüente com as redes IP.

Categorizar métodos e técnicas de ataque é, a meu ver, uma tarefa difícil e quase sempre destinada a produzir resultados incompletos ou imprecisos. O desenvolvimento contínuo de novas técnicas, o aperfeiçoamento e adaptação das técnicas existentes e a combinação de diversas técnicas em métodos de ataque mais complexos fazem com que toda classificação tenha sua precisão e abrangência rapidamente comprometida. Contudo, descrever as técnicas de ataque com base em suas características similares constitui-se numa forma útil de esclarecer seu funcionamento, expor os conceitos envolvidos em sua criação e avaliar os métodos de prevenção e reação que podem ser opostos a elas. A classificação proposta neste trabalho é simples e orientada ao objetivo genérico que cada uma das técnicas busca alcançar. Sendo assim, procuro organizar as técnicas de ataque com base nos efeitos pretendidos a partir de seu emprego, estabelecendo três grandes categorias :

- **Técnicas para obtenção de informações**
- **Técnicas para negação de serviços**
- **Técnicas para obtenção de acesso não autorizado**

Algumas técnicas, entretanto, acham seus aspectos mais marcantes não nos efeitos pretendidos, que podem ser bastante variados, mas na natureza do sistema ou sub-sistema a que se direcionam. Para melhor caracterizá-las foram criadas três categorias especiais :

- **Ataques contra roteadores**
- **Ataques baseados em CGI**
- **Ataques baseados em browsers**

1. <u>Técnicas para obtenção de informações</u> :

Reuni nessa categoria as técnicas e procedimentos que podem ser usados para a obtenção de informações críticas sobre a rede ou o host alvo. Tipicamente, essas informações serão usadas a posteriori na orientação de outros ataques. Conforme o critério de classificação proposto aqui, a principal finalidade dessas técnicas é obter informações sobre sistemas. O acesso a dados comunicados pela rede ou armazenados

em hosts conectados não constitui-se em seu objetivo prioritário, embora algumas delas possam fornecer isso, cumulativamente ou num segundo momento. Dentre as informações mais freqüentemente procuradas estão :

- **Nome do host alvo**
- **Endereço IP do host alvo**
- **Topologias física ou lógica da rede**
- **Sistema operacional executado no alvo**
- **Aplicações em execução no alvo**
- **Atualizações ou patches de correção instalados**
- **Serviços disponíveis no alvo**
- **Portas em estado de escuta (listening)**
- **Senhas de usuários ou administradores do sistema**

Algumas das técnicas mais conhecidas desta categoria, descritas nos sub-tópicos a seguir, são as seguintes :

- **Cracking de senhas**
- **Varredura de portas**
- **Emprego de utilitários**

1. **Cracking de senhas :**

   As senhas constituem-se em uma informação crítica sobre o sistema na medida em que podem permitir o acesso a funções, recursos e dados.
   O cracking ou quebra de senhas é, na verdade, um extenso conjunto de métodos e técnicas por meio dos quais procura-se recuperar as senhas que são usadas por usuários e administradores do sistema alvo a partir dos arquivos onde são registradas ou dos datagramas que as transportam ao longo da rede. Esse processo envolve, na maioria dos casos, as seguintes exigências :

   - **É necessário obter a senha em texto claro, de modo que possa ser reproduzida para utilização.**

- **É necessário que a senha obtida seja associada a uma identificação de usuário válida, uma vez que o processo de login exige a identificação do usuário juntamente com a informação da senha.**

- **Nos casos em que se procura obter a senha a partir de arquivos, é necessário obter acesso aos arquivos. A possibilidade de cópia desses arquivos para um ambiente controlado, isento do efeito de contramedidas, favorece o êxito das técnicas aplicadas.**

- **Nos casos em que se procura obter a senha a partir de datagramas interceptados durante seu trânsito pela rede, é necessário capturar e armazenar esses datagramas, conferindo-lhes a estabilidade necessária à aplicação da técnica escolhida.**

A fim de que essas exigências possam ser atendidas, as técnicas para quebra de senhas costumam ser empregadas em conjunto com outras técnicas, como o sniffing passivo, que será detalhado mais adiante, ainda no âmbito desta categoria.

As técnicas para quebra de senhas são desenvolvidas com base em dois métodos genéricos, que chamaremos de método de força bruta e de método do dicionário.

O método de força bruta é um método de tentativa e erro, envolvendo o teste de todas as combinações possíveis de caracteres que podem ser usados como senha, até que uma combinação funcione. Este método é pouco eficiente, em especial quando são usadas senhas longas ou quando as senhas são compostas com uma grande variedade de caracteres alfabéticos, numéricos e especiais.

O método do dicionário emprega uma lista previamente elaborada de palavras ou combinações de caracteres, que são testados um a um como possível senha, até que uma combinação produza um resultado positivo. Este

método reduz o espaço amostral das combinações possíveis, porém de forma arbitrária. É, em geral, mais rápido e eficiente que o método de força bruta, salvo nos casos em que as senhas são geradas aleatoriamente, não correspondendo a palavras de uso comum ou outras combinações conhecidas. Uma variante desse método consiste em testar como possíveis senhas certas palavras ou combinações que sejam, por alguma razão, relacionadas ao usuário cuja senha pretende-se obter. Um pequeno "dicionário pessoal" pode ser elaborado com o próprio nome do usuário, a data de seu nascimento, outra senha do usuário que já seja conhecida e informações similares, ajudando a reduzir o tempo necessário para obter-se um resultado positivo.

Esses métodos são implementados em uma grande variedade de ferramentas específicas para a quebra de senhas. O funcionamento dessas ferramentas é baseado num processo contínuo de geração e teste de alternativas, com um elevado custo em termos de tempo e capacidade de processamento. A vulnerabilidade de um sistema ao cracking de senhas não deve ser avaliada pela possibilidade ou impossibilidade em obter-se a senha correta. Teoricamente, qualquer senha pode ser quebrada. A avaliação deve levar em conta esse custo e os limites práticos que ele impõe ao atacante.

Todos os sistemas operacionais oferecem mecanismos para a utilização e gerenciamento seguros das senhas. Esses mecanismos podem apresentar níveis variados de sofisticação e eficiência, mas quase sempre envolvem criptografia em algum nível, controle sobre o tempo de validade de cada senha, bloqueio de acesso após tentativas de login frustradas e crítica e rejeição de senhas óbvias ou demasiado curtas, as chamadas senhas fracas. O mau uso desses recursos ou a excessiva confiança neles são apontados consensualmente como as principais vulnerabilidades dos sistemas a essa técnica. Administradores e usuários costumam utilizar ou permitir o uso de senhas demasiado curtas, que facilitam o trabalho de cracking por exigirem um número relativamente pequeno de tentativas para serem descobertas. A confiança no fato de que o sistema não permitirá ataques diretos, bloqueando o acesso a uma conta quando forem verificadas várias tentativas de login

incorretas, ou a idéia de que a criptografia aplicada às senhas é totalmente eficaz podem dar margem a uma proteção inadequada aos arquivos que as contém ou à negligência quanto à possibilidade de que senhas criptografadas possam ser capturadas na rede. Nesses casos, é deixado ao atacante o recurso de testar alternativas criptografando-as com algoritmos conhecidos e comparando o resultado com a senha criptografada a que teve acesso.

2. **Varredura de portas :**

A varredura de portas, ou port scanning, consiste em testar as portas de um host, ou mesmo de um grupo de hosts, a fim de determinar quais dessas portas estão em condições de aceitar conexões. Eventualmente, a varredura de portas é capaz de revelar outras informações de interesse, tais como o sistema operacional em execução, a partir da análise de como o alvo reage aos eventos gerados na varredura.

A varredura de portas é feita pelo uso de software específicos, capazes de enviar pacotes com destino às portas do host alvo, monitorar as respostas emitidas pelo alvo e gerar relatórios de análise.

Um software precursor dentre os port scanners é o SATAN, ou System Administrator Tool for Analyzing Networks. Ele foi criado por Dan Farmer e Wietse Venema como uma ferramenta de análise para uso de administradores de redes. SATAN é capaz de analisar um host específico ou um grupo de hosts indicados, procurando por ocorrências de uma extensa série de vulnerabilidades conhecidas, identificando-as e gerando relatórios detalhados. Ironicamente, essas funcionalidades podem ser tão úteis para a prevenção de ataques e falhas quanto para a identificação de alvos compensadores.
Podem ocorrer ligeiras diferenças na forma como o port scanner executa a varredura. Teríamos, assim, os seguintes tipos de varreduras de portas :

- **Varredura padrão – nesses casos a varredura é iniciada com o envio de um pacote TCP SYN para o alvo. O three-way handshake é executado**

de forma completa e a conexão entre o host que executa o port scanner e o alvo é estabelecida. Isso possibilita a obtenção de informações mais abrangentes, mas tem o inconveniente de facilitar que a varredura seja percebida e sua origem determinada.

- **Varredura TCP SYN – nesses casos a varredura também é iniciada com o envio de um pacote TCP SYN para o alvo, mas a conexão não é completada. Quando o alvo responde à solicitação de abertura da conexão com um pacote TCP ACK ou TCP RST, o port scanner já é capaz de avaliar se a porta está em estado de escuta e interrompe o processo de conexão com aquela porta. Por essa razão, essa modalidade é referenciada às vezes como varredura semi-aberta. O método dificulta a detecção da varredura.**

- **Stealth scanning – nessa modalidade o processo é iniciado não com um pacote TCP SYN, mas com um pacote que simula uma conexão já existente. A forma como o alvo responde a esse pacote pode não só revelar o estado da porta mas também fornecer informações adicionais sobre o alvo. Este tipo de varredura foi desenvolvido com o propósito de evitar que filtros de pacotes configurados para bloquear apenas pacotes TCP SYN inviabilizem a varredura.**

A varredura de portas é uma técnica de ataque que se aproveita do fato de que o TCP/IP gerencia as conexões de forma automatizada e com um nível de crítica muito elementar. Na maioria dos casos, um host procura responder a toda e qualquer solicitação de abertura de conexão endereçada a suas portas, sem avaliar a origem do pedido. O uso de portas conhecidas associadas a serviços padrão é outra característica explorada por essa técnica.

3. **Emprego de utilitários :**

Esta técnica apoia-se no emprego de ferramentas padrão de gerenciamento da rede com o objetivo de obter informações sobre o alvo. Se por um lado oferece

resultados mais limitados que as demais técnicas nesta categoria, por outro o emprego de utilitários é de execução simples e rápida, dispensando, em alguns casos, a utilização de software específico, uma vez que vários dos utilitários que dão suporte à técnica fazem parte dos conjuntos TCP/IP padrão.

Uma configuração cuidadosa de serviços nos hosts que representam alvos potenciais, associada a outras medidas preventivas, pode limitar muito a eficácia desta técnica. Contudo, a experiência comprova que o emprego de utilitários é ainda um método eficaz de obtenção de informações sobre sistemas e ambientes.

Dentre uma grande variedade de utilitários que podem ser empregados, citamos os seguintes :

- **Ping – é um utilitário utilizado para verificar se um determinado host está em condições de ser conectado. Ping funciona enviando uma mensagem ICMP_ECHO a um determinado host, especificado por seu nome ou pelo seu endereço IP. O host receberá a mensagem e responderá automaticamente com uma mensagem ICMP_REPLY, informando seu endereço IP ou seu nome e o tempo de ida e volta das mensagens. Ping é um dos utilitários mais simples dentre aqueles que encontramos associados ao TCP/IP e está disponível em praticamente todas as suas implementações.**

- **Traceroute – é um utilitário que pode ser usado para determinar a rota que um datagrama deverá percorrer para atingir um determinado host. Traceroute envia ao host de destino um datagrama com TTL igual a 1. Ao ser descartado pelo primeiro roteador no caminho, uma mensagem ICMP é gerada e endereçada ao host de origem, que repete o envio do datagrama incrementando o TTL até que o host de destino seja alcançado. Isso permite que todos os roteadores intermediários sejam identificados.**

- **Netstat e nbtstat – o netstat é um utilitário presente na maioria das implementações TCP/IP. O nbtstat é utilizado em implementações que utilizam o protocolo NetBIOS sobre o TCP/IP, como as implementações TCP/IP da Microsoft. Estes utilitários são capazes de fornecer informações sobre portas, interfaces de rede e tabelas de roteamento, entre outras.**

- **Finger – o finger constitui-se em outro exemplo de utilitário padrão do TCP/IP que pode ser utilizado para a obtenção de informações sobre o sistema alvo. A utilização do finger pode revelar detalhes como nomes de usuários conectados ao sistema. A utilização do finger depende de que o serviço correspondente esteja habilitado no host alvo.**

4. **Sniffing passivo :**

Esta técnica consiste na utilização de sniffers, que são componentes de software capazes de capturar os pacotes que trafegam por uma rede. Os sniffers incluem drivers que permitem que as interfaces de rede passem a operar em modo promíscuo, isto é, escutando todos os pacotes que circulam pelo meio físico ao qual estão conectadas e não apenas aqueles que são destinados a elas. Os sniffers podem filtrar os pacotes recebidos, capturando somente os que atenderem a determinados critérios de interesse pré-configurados. Após capturados, os pacotes podem ser armazenados para uma posterior análise de suas características ou das informações neles contidas.

O emprego de sniffing passivo exige algum tipo de acesso a um host da rede local, onde o sniffer será instalado. Por essa razão, a técnica é mais freqüentemente empregada por usuários internos da rede ou por usuários remotos após o prévio comprometimento de um host local.
Apesar dessas exigências, nem sempre fáceis de atender, sniffing passivo é uma técnica muito eficaz e seu emprego pode representar uma ameaça considerável à confidencialidade de informações críticas, como senhas por

exemplo. Na verdade, as possibilidades do sniffing passivo estendem-se além da obtenção de informações relativas ao funcionamento dos sistemas, criando condições para que seja obtido também acesso não autorizado aos dados que trafegam na rede.

O emprego de sniffers dá margem ainda a uma extensão da técnica com objetivos mais amplos que a simples obtenção de informações. Os pacotes capturados pelos sniffers podem vir a ser alterados e retransmitidos, num procedimento denominado sniffing ativo ou active sniffing.

As ferramentas empregadas em sniffing são variadas e de fácil obtenção. A maioria executa sobre UNIX ou Windows NT. Boa parte delas são de uso comercial, constituindo-se em módulos de software de gerenciamento ou de análise de tráfego de rede. O Network Monitor, desenvolvido pela Microsoft para uso com Windows NT, é um bom exemplo disso.

Sniffing é uma técnica de ataque que oferece elevado potencial de risco. Por isso, as contramedidas indicadas para prevenir ou controlar o seu emprego são altamente recomendadas, ainda que envolvam alguma complexidade. Dentre essas contramedidas, citamos as seguintes :

- **Criptografia de dados e de parâmetros – a aplicação de criptografia sobre os dados transmitidos na rede, bem como sobre parâmetros contidos nos cabeçalhos de camadas superiores, reduz o impacto desta técnica, dificultando o acesso do atacante às informações capturadas.**

- **Emprego de uma topologia segura – para capturar pacotes em trânsito, o sniffer exige mídia compartilhada, isto é, requer que o adaptador de rede colocado em modo promíscuo seja conectado diretamente a um segmento de rede. Nessas condições, os pacotes capturados serão exclusivamente os que trafegam por aquele segmento. A organização física da rede em segmentos distintos, conectados entre si por roteadores, switches ou bridges, irá contribuir**

para reduzir as amostras de dados capturados, limitando a eficácia desta técnica. Uma topologia segura deve também reduzir os riscos de que dispositivos estranhos sejam anexados à mídia.

- **Detecção de sniffers em atividade– a detecção de sniffers em uma rede é uma operação difícil. De início, os sniffers são passivos, limitando-se a ouvir o tráfego na rede sem emitir sinais de sua presença. Existem software desenvolvidos com a finalidade de detectar sniffers, mas seu emprego é complexo e limitado a determinados sistemas operacionais. A verificação em cada host pode ser uma alternativa mais eficiente, sobretudo em redes pequenas ou médias. Recursos de hardware ou software que detectem e informem alterações na topologia da rede, como a inclusão indevida de um novo host, podem contribuir circunstancialmente para a detecção de um sniffer instalado.**

2. <u>Técnicas para negação de serviços</u> :

Esta categoria engloba um conjunto de técnicas freqüentemente citadas sob a denominação genérica de DoS, do inglês Denial of Services. Seu propósito comum é o de tornar hosts e serviços da rede inoperantes ou, pelo menos, degradar seu funcionamento a um ponto em que sua utilização torne-se inviável ou não compensadora.

A maioria dessas técnicas busca esgotar os recursos do sistema alvo, forçando uma interrupção total ou parcial dos serviços. Capacidade de processamento, largura de banda e capacidade de armazenamento de dados são alguns dos recursos visados pelas técnicas de negação de serviços. Falhas de implementação, bugs e outras peculiaridades dos sistemas operacionais e dos aplicativos são também considerados no desenvolvimento dessas técnicas, na medida em que podem oferecer oportunidades para comprometer o funcionamento do sistema, a partir de sua incapacidade de tratar convenientemente erros ou eventos específicos.
Embora existam técnicas de negação de serviços destinadas a atacar estações de trabalho e computadores pessoais, como o célebre WinNuke, que explorava falhas de implementação nas primeiras versões do Windows 95 para causar quedas de conexão e

travamentos, a maioria dos ataques parece ser orientada contra sistemas de maior porte, que oferecem serviços a um grande número de usuários, como é o caso de servidores web, ou que desempenham funções críticas para o funcionamento de redes e sistemas distribuídos. Roteadores, servidores DNS e firewalls costumam ser os alvos preferenciais desses ataques.

É interessante notar que as técnicas de negação de serviços são freqüentemente aplicadas como uma etapa intermediária de métodos de ataque mais complexos, servindo, por exemplo, para deixar um host fora do ar a fim de que outro host assuma sua identidade ou interrompendo o funcionamento de um sistema que execute funções de segurança e controle da rede.

Serão citadas as seguintes técnicas de negação de serviços :

- **SYN Flood**
- **LAND**
- **Ataques baseados em ICMP**
- **Teardrop**
- **Ping o'Death**
- **Ataques de dessincronização**
- **DDoS – Distributed Denial of Services**

1. **SYN Flood :**

   Esta técnica consiste no envio de uma grande seqüência de pacotes TCP SYN para o host alvo. Esses pacotes são interpretados como solicitações de estabelecimento de conexão e, como tais, são armazenadas em uma fila de conexões em andamento. Os pacotes TCP SYN/ACK enviados pelo alvo em resposta aos falsos pedidos de conexão não são respondidos, normalmente porque as solicitações são geradas com um endereço IP falso ou inválido no lugar do endereço da máquina que originou o ataque. Assim, a fila de conexões em andamento atinge rapidamente seu limite configurado e passa a

descartar novas solicitações de conexão, o que na prática torna indisponíveis os serviços residentes no host alvo.

É claro que o alvo não irá manter os pedidos de conexão em andamento indefinidamente. Ele os eliminará num intervalo de tempo pré-definido. Contudo, enquanto o envio de falsas solicitações de conexão for contínuo, a efetividade do ataque será mantida.

A vulnerabilidade de um sistema a este tipo de ataque pode ser reduzida se forem adotadas configurações especiais que aumentem a capacidade da fila, permitindo que um número maior de conexões permaneçam aguardando, ou que reduzam o limite de tempo após o qual uma conexão solicitada e não estabelecida é eliminada. O emprego de software específico para monitorar os pedidos de conexão em andamento constitui-se em uma contramedida ainda mais eficaz.

2. **LAND :**

LAND baseia-se no efeito que o recebimento de um datagrama IP no qual os endereços de origem e destino são iguais pode ter sobre determinados sistemas. Em certos casos, o processamento desse datagrama no alvo irá produzir um loop que pode esgotar os recursos da máquina, causando sensível queda de desempenho ou até o travamento do sistema.

Variações desta técnica prevêem alterações em outros campos no cabeçalho IP do datagrama inválido, como as portas ou os bits de controle, por exemplo. Essas alterações são capazes de produzir no host alvo efeitos semelhantes aos do LAND original.

Os datagramas inválidos são enviados pelo atacante mediante o uso de software específico que permite a geração do datagrama e a adulteração dos campos conforme sua necessidade.

3. Ataques baseados em ICMP :

Este conjunto de técnicas aproveita-se das funcionalidades do ICMP – Internet Control Message Protocol, para criar eventos capazes de afetar o funcionamento de alvos específicos. Como vimos, o ICMP é empregado em tarefas de controle e verificação das comunicações entre hosts e roteadores. Ele utiliza mensagens padronizadas que são enviadas com o propósito de checar a possibilidade de comunicar-se com um host de destino. Por exemplo, o utilitário ping, bastante conhecido, permite o envio de mensagens ICMP a um host especificado a fim de verificar se esse host é alcançável. As implementações padrão do TCP/IP irão reagir automaticamente às mensagens ICMP recebidas, executando as ações apropriadas a cada caso. Elas responderão a pedidos de ICMP_ECHO ou poderão encerrar conexões estabelecidas a partir do recebimento de mensagens do tipo Destination Unreachable ou Time to Live Exceeded.

Normalmente, não será executada qualquer modalidade de autenticação dos pedidos ou de crítica de características especiais, como repetições excessivas.

Uma das modalidades mais simples de ataque baseado em ICMP é o PingFlood. Nessa modalidade uma seqüência ininterrupta de mensagens ICMP é enviada ao host alvo, que ocupa-se em responder a todas elas, consumindo desnecessariamente seus recursos.

Por sua vez, o ataque denominado Pong envia mensagens ICMP a um grande número de hosts, normalmente endereçando-as a um endereço de broadcast. Nas mensagens, o endereço de resposta informado é o endereço do alvo. Quando todos os hosts respondem, o alvo recebe uma grande quantidade de mensagens ICMP simultaneamente, tendo com isso suas comunicações afetadas ou interrompidas.

A técnica denominada SMURF constitui-se num aperfeiçoamento do Pong, ampliando o número de hosts que enviarão mensagens ICMP ao alvo pelo

envio da requisição falsa não apenas a um mas a vários endereços de broadcast. Uma variação dessa técnica, que foi denominada FRAGGLE, utiliza como protocolo de transporte o UDP em lugar do TCP.

4. **Teardrop :**

   Esta técnica explora o processo de remontagem de um datagrama IP fragmentado, adulterando informações no cabeçalho IP de forma a produzir uma situação que o alvo não consiga processar adequadamente, causando instabilidades ou falhas.

5. **Ping O'Death :**

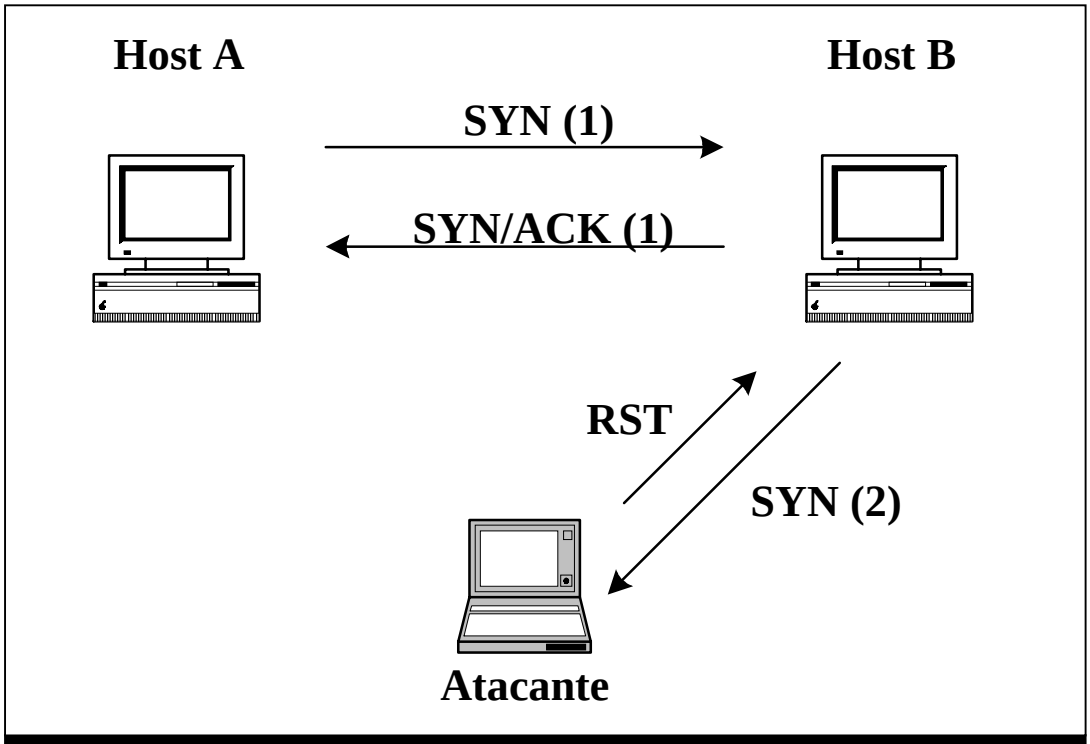
   Apesar do nome, o Ping O'Death não é um ataque baseado em ICMP, embora suas primeiras versões tenham sido baseadas em uma modificação do utilitário ping capaz de gerar os datagramas inválidos empregados neste ataque. Na verdade, a técnica consiste basicamente no envio ao host alvo de um datagrama com tamanho além do limite de 65535 bytes. Como um datagrama com essas características não pode ser roteado, ele é fragmentado na origem e chega ao alvo sob a forma de vários datagramas contendo os fragmentos do datagrama original. Quando o alvo procede a remontagem do datagrama a partir desses fragmentos, ele experimenta efeitos que podem variar, conforme o sistema operacional em uso, da exibição de mensagens de erro até a completa paralisação do sistema.

6. **Ataque de dessincronização :**

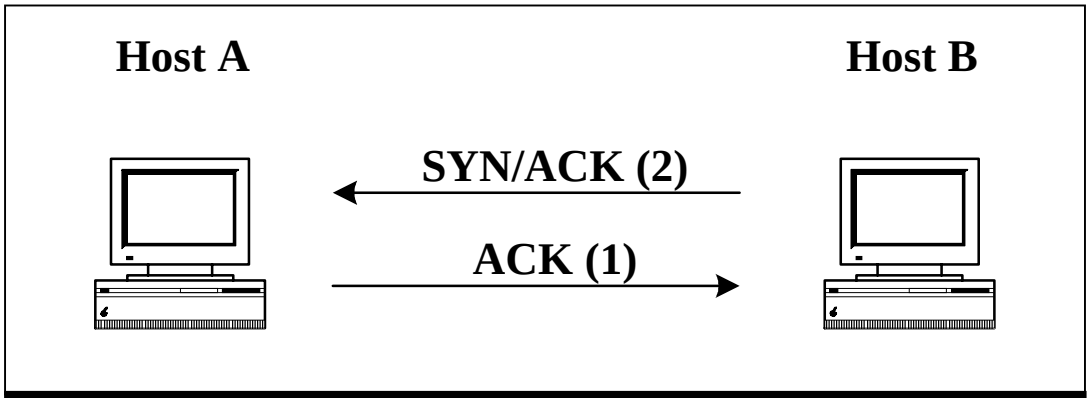
   A interferência nos estágios iniciais da conexão TCP é a base desta técnica. O atacante monitora o envio de pedidos de abertura de conexão de um host a outro. Quando o host B reconhece o pedido de abertura de conexão, enviando ao host A um segmento SYN/ACK, o atacante envia-lhe um segmento RST com o endereço IP e a porta utilizados pelo host A. Isso leva o host B a finalizar prematuramente a conexão. Imediatamente, o atacante envia ao host

B um segmento SYN, solicitando a abertura de uma nova conexão. Esse segmento também leva o endereço IP e a porta utilizados pelo host A. O host B aceita o pedido e abre uma segunda conexão. Como o host A não tem conhecimento da finalização da primeira conexão provocada pelo atacante, ele mantém a porta utilizada em estado de conexão estabelecida. O host B também manterá sua porta correspondente em estado de conexão estabelecida, pois acredita que a nova conexão estabelecida com o host A é válida. O resultado desse processo é que os dois hosts terão estabelecido uma conexão dessincronizada, que não pode ser usada para a transferência de dados mas que consome recursos de ambas as máquinas.

As figuras abaixo ilustram as duas fases desse processo :



Ataque de dessincronização (primeira fase).



Ataque de dessincronização (segunda fase).

7. DDoS – Distributed Denial of Services :

Esta técnica ganhou grande notoriedade a partir da série de ataques bem sucedidos empreendidos contra sites famosos no início do ano 2000. Ela potencializa os danos causados por outras técnicas de negação de serviços, como SYN Flood por exemplo, fazendo com que um grande número de hosts lancem o ataque, de forma automatizada e simultânea, contra um alvo escolhido.

O ataque deve ser previamente organizado e para isso são empregados software específicos como o TFN, o Stacheldraht e o Trinoo. Essas ferramentas são instaladas em alguns hosts, deixando-os em condições de atuar como servidores ou masters. Paralelamente, um grande número de outros hosts recebe também componentes de software, passando por sua vez a representar o papel de clientes ou slaves. Tanto a instalação de servidores quanto a distribuição dos módulos clientes são feitas de forma não autorizada, com os componentes embutidos em outros programas supostamente inofensivos ou aproveitando-se de conhecidas falhas de segurança dos sistemas operacionais.

A comando do atacante, os servidores comunicam-se com os clientes, determinando o início do ataque. Em resposta, os hosts que executam o módulo cliente lançam, a um só tempo, uma série de ataques contra o alvo ou os alvos especificados no comando inicial. A efetividade dos ataques é aumentada pela participação simultânea de um elevado número de hosts. O atacante goza da vantagem adicional representada pela dificuldade em determinar a origem do ataque, que em sua etapa final é desencadeado a partir de diversos pontos, normalmente sem a intervenção dos responsáveis pelos hosts que o executam.

As ferramentas utilizadas em DDoS apresentam um alto nível de sofisticação, integrando recursos avançados que vão desde mecanismos de distribuição automatizada dos módulos clientes até comunicações criptografadas entre os servidores e os clientes.

Distributed Denial of Service merece atenção não apenas pela sua eficácia, mas principalmente por estabelecer um novo modelo de ataque distribuído. Entretanto, as versões conhecidas do ataque exploram as mesmas brechas exploradas por outras técnicas de negação de serviços, não tendo contribuído até o momento para revelar novas falhas de segurança, mas apenas novos graus de comprometimento a partir de vulnerabilidades conhecidas.

3. <u>Técnicas para obtenção de acesso não autorizado</u> :

A característica comum entre as técnicas de ataque agrupadas nesta categoria é a de terem como propósito principal oferecer ao atacante acesso não autorizado a dados ou serviços dos sistemas alvo. Uma vez mais, os efeitos obtidos a partir da aplicação destas técnicas podem mesclar-se. Ao utilizar uma delas para transferir dados sem autorização a partir de um sistema remoto, um atacante pode, eventualmente, obter acesso a informações críticas do alvo, como um arquivo de senhas, por exemplo. Numa outra situação igualmente plausível, se um atacante assume o endereço IP de um determinado host a fim de obter para si o acesso a recursos normalmente concedido a esse host, os serviços fornecidos pelo host personificado ficarão inacessíveis enquanto durar o ataque.

Deve-se ter em mente que o acesso não autorizado pode envolver riscos mais amplos do que simplesmente fornecer conhecimento sobre a informação representada nos dados. Dependendo do grau de acesso obtido, o atacante pode ver-se em condições de alterar o conteúdo dos dados sem que isso seja percebido. Os dados adulterados podem ser reenviados ao destino ou mantidos na sua localização original sem alterações aparentes.

Os efeitos dessa ameaça potencial podem ser, em determinadas circunstâncias, ainda mais agudos que a simples quebra da confidencialidade dos dados.

Serão descritas a seguir as seguintes técnicas de ataque :

- **Spoofing**
- **Active sniffing**
- **Buffer overflow**
- **Exploração de relacionamentos de confiança**
- **Ataques de fragmentação**
- **Ataques baseados em sequence numbers**
- **Trojans**

1. **Spoofing**

   O spoofing é o ataque de personificação típico, consistindo na utilização de um endereço IP falso por parte do atacante, normalmente com o propósito de simular ou personificar outro host. Isso pode ser feito mediante alterações no cabeçalho IP dos pacotes transmitidos ou, quando possível, configurando com um endereço IP falso o host de onde será lançado o ataque.

   O spoofing é usado freqüentemente para explorar uma relação de confiança ou um mecanismo de autenticação baseados unicamente em endereço IP. De fato, existem vários serviços que concedem acesso a endereços IP especificados, como o NFS e o serviços de proxy do MS Windows NT.

O spoofing terá sua execução mais fácil se o atacante puder abrir a conexão usando o IP falso, em lugar de inserir-se em uma conexão já existente. Na verdade, as técnicas que envolvem a personificação de um host no curso de uma conexão estabelecida apresentam diferenciais de complexidade tão apreciáveis que são consideradas técnicas à parte, como terei oportunidade de expor mais adiante. Para que o atacante encontre condições de abrir a conexão simulando ser um outro host, é necessário que o outro host não esteja ativo no momento. Isso pode ser conseguido aproveitando uma ocasião em que o host esteja efetivamente desligado, o que pode ser verificado ou mesmo previsto por uma criteriosa coleta de informações sobre os alvos. Caso o host que será personificado ache-se em atividade, a aplicação de uma técnica de negação de serviços pode tirá-lo do ar, fornecendo ao atacante a oportunidade de assumir seu endereço IP.

A prática do spoofing é simplificada quando aplicada sobre UDP. Nesse caso, a simplicidade relativa do protocolo favorece o emprego da técnica. Em contrapartida, em conexões TCP o atacante estaria às voltas com o problema representado pelos sequence numbers que acompanham cada pacote transmitido e que precisariam também ser corretamente simulados, sob pena de que a conexão fosse interrompida.

Uma situação que seria extremamente favorável ao atacante é aquela na qual o host de onde o ataque é lançado encontra-se na mesma rede local que o host a ser personificado. Nesses casos, roteadores e firewalls não seriam capazes de perceber a simulação, uma vez que o endereço IP é coerente com a localização do host. Porém, se o atacante tenta simular um endereço IP a partir de uma rede remota, o que representa uma situação mais usual, um firewall bem configurado poderá perceber que os pacotes transmitidos pelo atacante não se originam da rede local ou interna como deveriam. Espera-se, nessas circunstâncias, que o firewall filtre os pacotes e anule o ataque.

Outro aspecto a ser considerado é que o emprego desta técnica poderá ter algum impacto sobre a segurança dos sistemas ainda que o atacante não se coloque em condições de receber os pacotes enviados pelos hosts iludidos em resposta às suas solicitações. Isso dependerá do propósito do ataque. Caso esse propósito esteja limitado a obter acesso para acrescentar dados ou arquivos, como um cavalo de tróia por exemplo, no sistema alvo, pouco importará se os pacotes de resposta se perdem ou são enviados ao host personificado.

Um caso particular de spoofing digno de menção é o spoofing por meio de DNS. Nesse caso, o atacante altera a base de dados um servidor DNS de modo a associar seu endereço IP a uma determinada URL. As requisições feitas àquele servidor envolvendo a URL alterada serão resolvidas com o fornecimento do endereço IP do atacante. Isso irá direcionar o acesso de clientes ao host do atacante em lugar do host ou serviço pretendido. Se o host do atacante oferecer um serviço similar ao solicitado pelo cliente, como uma página web forjada por exemplo, é possível que o engano nem seja percebido e que o cliente forneça informações ou recupere dados não íntegros ou não autênticos.

2. Active sniffing

Active sniffing é uma variação da técnica passive sniffing, mencionada anteriormente, que consiste na captura de pacotes em trânsito pela rede. O active sniffing ocorrerá quando o atacante, além de capturar os pacotes, puder alterar seu conteúdo e reenviá-los, comprometendo sua integridade ou confidencialidade sem interromper as comunicações. Os métodos e ferramentas usados em active sniffing são, de maneira geral, os mesmos que suportam o passive sniffing. Da mesma forma, as contramedidas indicadas para prevenir a aplicação com êxito de uma técnica serão igualmente úteis contra a outra.

3. Buffer overflow

Esta denominação designa um conjunto de ações e métodos destinados a explorar uma classe de vulnerabilidades que afeta grande número de sistemas operacionais e aplicativos. Embora ligeiramente diversificados, esses métodos compartilham um princípio comum e buscam um único objetivo genérico.

O princípio comum baseia-se no fato de que alguns processos executados em um host são, usualmente, controlados por contas de sistema. Essas contas possuem direitos muito amplos, normalmente equivalentes aos direitos de administrador ou root. É o caso da conta SYSTEM no MS Windows NT ou da conta HTTPD em UNIX, para citar apenas dois exemplos. Compreendendo essa característica é possível imaginar que a execução de comandos arbitrários no âmbito de um processo originalmente conduzido por uma conta de sistema poderá ter efeitos críticos, uma vez que não estará sujeita às restrições de acesso impostas aos usuários comuns. O problema consiste em criar uma condição especial na qual os comandos inseridos pelo atacante sejam executados como parte do processo afetado. Essa condição é criada pelo chamado buffer overflow, ou estouro de buffer. O buffer overflow é desencadeado pela entrada de dados com extensão ou formato que o processo não seja capaz de tratar, provocando uma sobrecarga no espaço de memória alocado para a aplicação. Essa sobrecarga irá possibilitar a execução subsequente de código arbitrário, a partir de comandos que o atacante poderá inserir remotamente. Neste ponto, conforme a aplicação ou sistema visado, o atacante estará apto a acessar arquivos, obter um shell ou alterar configurações.

O objetivo genérico dos ataques de buffer overflow é obter acesso privilegiado a um sistema, dando ao atacante a possibilidade de executar remotamente tarefas que, em condições normais, seriam restritas.

O buffer overflow decorre de falhas na implementação de software. A ausência de rotinas de crítica aos dados de entrada é uma das mais comuns. A

identificação de vulnerabilidades que exponham um software específico a esta técnica de ataque costuma levar ao desenvolvimento e à distribuição gratuita de patches e atualizações. A aplicação criteriosa dessas correções é uma contramedida simples mas eficaz para o tratamento do problema.

4. Exploração de relacionamentos de confiança

Relacionamentos de confiança são um recurso disponível em sistemas como UNIX e MS Windows NT que visa estender os direitos de acesso concedidos a um usuário do sistema a outros sistemas, simplificando e flexibilizando a administração. No MS Windows NT os relacionamentos de confiança são estabelecidos entre domínios. No UNIX, podem ser estabelecidos entre hosts. Em qualquer caso, a essência do recurso é a de que um host ou domínio declarado ou configurado como confiável passa a ter acesso automático a recursos do domínio ou host confiante, normalmente sem que se faça necessário realizar autenticações ou validações adicionais.

Em que pese sua utilidade, mais sensível em ambientes de rede complexos, os relacionamentos de confiança representam uma ameaça potencial à segurança. Sua exploração pode ser feita a partir do emprego de técnicas como o spoofing, de modo que o atacante venha a personificar um host confiável para obter acesso não autorizado. Existem, contudo, outras formas de exploração desse recurso. No caso específico do UNIX, os arquivos /etc/hosts.equiv e ~/.rhosts são usados para listar os hosts e usuários considerados confiáveis. Um atacante que venha a ter acesso a esses arquivos pode incluir entradas de modo a atribuir o status de usuário confiável a si próprio, ou de host confiável à máquina que lhe for conveniente.

5. Ataques de fragmentação

Os ataques de fragmentação são conduzidos para explorar determinadas limitações das funções de filtragem de pacotes que são implementadas em roteadores e outros dispositivos da rede. Essas funções, em muitos casos,

verificam apenas pacotes de abertura de conexão. Se o primeiro fragmento de um pacote recebido não contiver um bit SYN definido, o filtro assume que o pacote não é um pacote de abertura de conexão e permite sua passagem.

O ataque consiste na manipulação e envio de datagramas fragmentados de forma que o bit SYN esteja definido não no primeiro fragmento, mas num dos fragmentos subsequentes. Espera-se com isso que o filtro de pacotes aceite todos os fragmentos, permitindo que eles alcancem o alvo, normalmente um host de destino na rede interna. No alvo, o datagrama será remontado e, eventualmente, o pedido de conexão será reconhecido.

6. Ataques baseados em sequence numbers

Os ataques baseados em sequence numbers são técnicas de grande complexidade que permitem ao atacante forçar sua participação em uma conexão personificando um dos hosts envolvidos. O fundamento desta técnica é a previsão de sequence numbers.

Os sequence numbers, como vimos no capítulo 1, são parâmetros usados para manter a conexão baseada em TCP, sendo incrementados a partir de um valor inicial, denominado ISN – Initial Sequence Number - e transmitidos por um host ao outro a cada mensagem enviada ou recebida. Os sequence numbers orientam a reordenação dos segmentos que chegam ao host de destino e seu correto recebimento sinaliza para cada host que seu interlocutor continua participando da conexão.

A cada pedido de conexão recebido é informado um ISN. As regras para a geração do ISN envolvem uma variável aleatória e um incremento fixo e padronizado, que será aplicado em uma frequência determinada sobre essa variável para definir cada novo ISN gerado. Como as características do incremento padrão podem ser conhecidas, a previsão do ISN é feita a partir de um ou mais pedidos de conexão ao alvo e da observação dos ISNs informados.

Como descrito por Bellovin em *Security Problems in the TCP/IP Protocol Suite*, fatores como o incremento padrão utiizado, o clock da CPU do alvo e mesmo a velocidade da rede podem contribuir para a o acerto da previsão. Curiosamente, CPUs e redes mais rápidas tendem a tornar as previsões mais apuradas.

Duas classes de contramedidas são recomendáveis na prevenção de ataques baseados em sequence numbers. A primeira diz respeito a alterações no incremento padrão, adotando incrementos menores ou frequências de aplicação maiores. Valores aleatórios para o incremento são também considerados. Naturalmente isso exige que as implementações do TCP sejam alteradas.

Uma contramedida mais ao alcance do administrador consiste em observar e registrar a ocorrência de pedidos de conexão não completados. Esses pedidos podem indicar a fase inicial de um ataque.

7. Trojans

Cavalos de Tróia, Trojan horses, ou simplesmente Trojans são programas inseridos em um sistema para executar automaticamente funções diversas, como captura de dados, alterações de características e configurações ou suporte a acesso remoto numa base cliente servidor. Trojans podem afetar redes de qualquer tipo, não estando vinculados a vulnerabilidades específicas dos protocolos TCP/IP. Contudo, devido à ampla utilização de TCP/IP em redes corporativas e sobretudo na Internet, a maioria dos Trojans que oferecem funções de comunicação baseia-nas em TCP/IP.

Trojans são instalados no sistema alvo por métodos muito variados. Tipicamente estão ocultos em meio ao código de outro programa ou são inseridos por meio de acesso não autorizado. No primeiro caso, reproduzem

alguns dos métodos de inserção ou "contágio" empregados pelos vírus de computador.

O uso de software específico para detectar e remover Trojans pode ser uma medida eficiente contra variantes cujo código seja conhecido. Controles de acesso eficientes e padronização de ambientes e de configurações são contramedidas menos focalizadas, mas que podem fornecer uma proteção mais abrangente.

4. <u>Ataques contra roteadores</u> :

Como vimos anteriormente, as funções de roteamento desempenham uma missão crítica em ambientes de redes interconectadas. Sob essa perspectiva, os ataques dirigidos contra os roteadores assumem enorme importância. Afetando os serviços de roteamento, podem comprometer a comunicação entre computadores de várias redes simultaneamente. Adulterando informações relativas à localização de redes e hosts, podem servir de ponto de partida para ataques de personificação e como medida de apoio a ataques de outros tipos.

Tratamos neste tópico das técnicas que têm como alvos específicos os roteadores e outros dispositivos de roteamento. Cumpre notar, todavia, que os roteadores podem ser vulneráveis a outras técnicas de ataque de aplicação genérica, como técnicas de cracking de senhas ou de negação de serviços. Isso é ainda mais provável nos casos em que o roteador é implementado sobre uma plataforma não-dedicada, como um servidor UNIX ou Windows NT. Nesses casos, além das vulnerabilidades específicas das funções de roteamento, precisam ser consideradas aquelas relacionadas ao sistema operacional em uso.

Algumas das formas mais características de ataque contra roteadores são as seguintes :

- **Ataques baseados em roteamento de origem.**
- **Ataques baseados no protocolo RIP**

- **Ataques baseados nos protocolos de roteamento externo**
- **Ataques baseados em ICMP**
- **Ataques baseados em SNMP**

1. **Ataques baseados em roteamento de origem :**

   Essa técnica abrange ataques contra dispositivos que empregam o roteamento de origem. Segundo esse método, um roteador procurará encaminhar os pacotes destinados a um determinado host utilizando a mesma rota percorrida pelos pacotes recebidos daquele host. Embora a idéia seja bastante coerente, na medida em que garante a transmissão de pacotes por rotas funcionais, ela implica numa falha de segurança evidente e de difícil controle. Um atacante que envie mensagens usando como IP de origem o endereço IP de outra máquina passaria a receber os pacotes destinados àquele host, ganhando acesso às informações que eles transportam. Nos casos em que o host personificado tenha recebido direitos de acesso com base em seu endereço IP, o atacante estaria em condições de usar desses mesmos direitos.

   A contramedida mais eficaz contra essa técnica de ataque consiste no emprego de firewalls. Os firewalls devem ser configurados para detectar e filtrar os pacotes originados de redes externas que contenham endereços de origem pertencentes à rede interna, evento que indicaria a tentativa de um host remoto de se fazer passar por um host local a fim de obter seus direitos e privilégios. Entretanto, em situações que envolvam relacionamentos de confiança estabelecidos entre redes diferentes, pode ser necessária a aplicação de procedimentos mais complexos, como será discutido no capítulo 4, que trata das contramedidas e de seu emprego.

2. **Ataques baseados no protocolo RIP :**

   Os pacotes RIP são trocados entre os roteadores para o compartilhamento e sincronização das tabelas de roteamento. A primeira versão do protocolo RIP

não suporta qualquer modalidade de autenticação. A versão 2, assim como o protocolo OSPF, permite o uso de procedimentos elementares de autenticação.

Os ataques baseados em RIP consistem no emprego de pacotes RIP forjados com o intuito de passar informações adulteradas aos roteadores, induzindo-os a encaminhar dados a endereços de destino falsos, conforme a conveniência do atacante. O endereço de destino falso pode ser de uma rede ou de um host isolado, embora o segundo caso seja mais comum, pelo fato de que a adulteração de informações de roteamento pode ser mais facilmente detectada quando se referem a redes inteiras.

A eficácia dessa técnica depende de que RIPv1 esteja habilitado nos roteadores. O emprego alternativo de RIPv2 ou OSPF, especialmente nos roteadores de fronteira, e a utilização dos mecanismos de autenticação oferecidos por esses protocolos representam uma contramedida eficiente, desde que o roteador que se autentica e fornece a informação não tenha sofrido, por sua vez, algum comprometimento. Na verdade, existe sempre o risco de que o efeito dos ataques baseados em protocolos de roteamento possa ser ampliado se as informações falsas forem propagadas a partir de um roteador comprometido para um grande número de dispositivos de roteamento.

Quando possível, a implementação nos roteadores de funções de crítica das informações relativas às redes diretamente conectadas também pode contribuir para evitar ataques desse tipo.

3. Ataques baseados nos protocolos de roteamento externo :

Os ataques baseados em protocolos de roteamento externo buscam, em geral, efeitos similares aos obtidos pelo emprego de outras técnicas de ataque contra o roteamento. A inserção de informações de roteamento falsas com o

propósito de personificar um host ou uma rede, causando com isso um redirecionamento do tráfego destinado a eles, é seu objetivo mais comum. Contudo, a implementação desses ataques é consideravelmente mais difícil, principalmente porque os protocolos de roteamento externo utilizam o TCP como protocolo de transporte. O tráfego de mensagens entre os roteadores é feito, portanto, por meio de conexões confiáveis e de mensagens sequenciadas. Além disso, existe uma rotina de comunicação que restringe o envio de mensagens não solicitadas entre os roteadores de sistemas autônomos diferentes. Normalmente, esses roteadores apenas respondem a solicitações de outros roteadores externos, não encaminhando espontaneamente as atualizações de suas tabelas ou, pelo menos, encaminhando-as somente em momentos bem definidos. Isso limita as oportunidades dadas ao intruso de divulgar uma informação de roteamento falsa, uma vez que elas dependerão de solicitações dos roteadores alvo. É possível, entretanto, que o atacante consiga personificar um roteador externo, talvez após interromper seus serviços mediante um ataque DoS. Isso o deixaria em condições de receber as solicitações enviadas por roteadores de outros sistemas autônomos e responder-lhes com informações adulteradas.

4. **Ataques baseados em ICMP :**

O Internet Control Message Protocol é usado por roteadores para divulgar rotas mais eficientes em um determinado momento. Através de mensagens Redirect o encaminhamento de pacotes pode ser corrigido de modo a utilizar uma rota alternativa com melhores resultados. Naturalmente, essa funcionalidade pode ser explorada por meio de mensagens Redirect geradas com o propósito de desviar o tráfego roteado para um novo itinerário. Isso só será possível, todavia, no âmbito de uma conexão em particular. Outra limitação desta técnica decorre do fato de que os roteadores só enviam mensagens Redirect para hosts localizados nas redes diretamente conectadas a eles.

5. Ataques baseados em SNMP :

Os dispositivos de roteamento implementam funções de gerenciamento e controle que permitem aos administradores obter informações sobre seu funcionamento e executar reconfigurações por meio de acesso remoto. O protocolo SNMP dá suporte à comunicação entre os agentes e as consoles de gerenciamento,

A estrutura de gerenciamento estabelecida pelo SNMP pode ser usada como base para ataques contra roteadores e outros dispositivos da rede. Aproveitando-se de contas padrão que foram deixadas habilitadas ou de práticas de autenticação pouco seguras, um atacante pode obter informações sobre a rede tomada como alvo e empregá-las em outras técnicas de ataque.

A prevenção desse comprometimento deve incluir, antes de tudo, o cancelamento de contas padrão nos dispositivos da rede e o emprego de autenticação segura entre as consoles e os agentes SNMP. A organização das consoles e agentes em comunidades específicas é recomendada. O monitoramento de tentativas de acesso mal sucedidas às bases de dados dos agentes complementa essas medidas, podendo fornecer indícios consistentes de que um ataque está sendo preparado.

5. <u>Ataques baseados em CGI</u>

Os ataques baseados em CGI utilizam protocolos da camada de aplicação e não estão vinculados diretamente aos protocolos IP, TCP ou UDP. Contudo, como as aplicações que dão margem a esses ataques são amplamente utilizadas nos ambientes típicos de redes IP, como a Internet ou as intranets corporativas, optei por mencioná-los aqui.

Os programas CGI são empregados com finalidades variadas, mas sobretudo para permitir que usuários remotos possam enviar ou recuperar dados de um computador remoto, tipicamente um servidor web, usando como interface uma

página exibida em browser. Em termos simplificados, o servidor web recebe os comandos enviados pelo browser e executa um programa CGI para processar os dados recebidos ou solicitados.

A técnica de ataque baseado em CGI engloba uma variedade de procedimentos que exploram a possibilidade oferecida pelos programas CGI para que seja executado código em um servidor a partir de comandos enviados de um host remoto. Um programa executado no servidor pode ser usado para produzir efeitos nocivos, dentre os quais a sobrecarga de processamento é o mais provável e fácil de obter. Outros objetivos, contudo, podem ser alcançados, como a interferência nas comunicações ou o acesso não autorizado a dados e informações do sistema. Essa possibilidade, longe de ser meramente teórica, pode concretizar-se pelo aproveitamento de recursos das linguagens de programação empregadas na codificação dos programas CGI. Scripts feitos em Perl, por exemplo, podem permitir que um atacante passe ao servidor comandos a serem executados em lugar de strings de dados. Ainda que exija conhecimento técnico e envolva considerável complexidade, o risco deve ser considerado.

Se for dado ao atacante a possibilidade de incluir um programa no servidor e executá-lo em seguida, esse risco é consideravelmente aumentado, uma vez que nesse caso o atacante contará com a vantagem adicional de poder utilizar o programa que for mais conveniente ao seu propósito. Por isso, é necessário que seja mantido um controle rígido sobre os programas disponíveis para execução no servidor, sobre sua localização na estrutura de arquivos, sobre a adição de novos programas ao conjunto e sobre que usuários podem executá-los.

6. <u>Ataques contra navegadores</u> :

Com o propósito de acrescentar funcionalidades mais amplas às páginas web, os modernos navegadores ou browsers vêm sendo dotados da possibilidade de executar código embutido em componentes das páginas, como scripts e applets Java ou Active X. Essa possibilidade dá margem ao desenvolvimento de todo um conjunto de

técnicas que podem ser usadas para negação de serviços, para o levantamento de informações do sistema alvo ou para obter acesso não autorizado a dados.

A execução de código a partir de componentes de páginas web delineia um problema de segurança. Em tese, um componente recebido como conteúdo de uma página web não deve ser capaz de executar determinadas ações no host cliente. Gravar no sistema de arquivos e registrar informações de configuração do host são algumas delas. É desejável ainda que os applets não possam provocar sobrecarga na máquina gerando processos que consumam demasiados recursos. Para atender a essas restrições, os applets ou miniaplicativos estão sujeitos a dois tipos de mecanismos de segurança : os controles impostos pela própria linguagem e as configurações de segurança do browser. Discutiremos brevemente cada um deles.

Em se tratando de controles impostos pela linguagem, podemos considerar Java como sendo mais seguro que Active X. De fato, Java dispõe de recursos como a Sandbox, um ambiente de execução isolado e gerenciado por uma classe Java específica, denominada SecurityManager, que restringe severamente as ações que o applet pode executar. Já o Active X contrapõe à sua versatilidade um nível de controle bem inferior e encerra um potencial para exploração que muitos consideram alarmante.

Funcionando nesse caso como uma segunda linha de defesa, os browsers dispõe de recursos para limitar as ações que podem ser executadas por miniaplicativos. Essas configurações devem ser ajustadas pelo usuário, considerando tanto quanto possível o grau de confiabilidade das páginas por ele visitadas. Não sendo possível determinar esse grau, o uso de configurações restritivas é recomendado.

Outro problema que afeta os browsers diz respeito aos plug-ins. Os plug-ins são componentes de software adicionados aos browsers para a execução de tarefas especiais, como a exibição do conteúdo de determinados tipos de arquivo. Usualmente os plug-ins são baixados e instalados automaticamnete durante a navegação, quando se fazem necessários. Há, portanto, o risco potencial de que um

plug-in seja adulterado de modo a incluir Trojans ou vírus e permitir a sua instalação na máquina cliente de forma inadvertida.

Os melhores programas anti-vírus são capazes de detectar e impedir a ação de boa parte dos componentes de páginas web que contém código malicioso, sejam eles miniaplicativos hostis ou plug-ins que tenham sofrido adulteração e incluam Trojans ou vírus.

Capítulo 4

Contramedidas

Contramedidas

Este capítulo apresenta algumas contramedidas que podem ser empregadas para evitar ou controlar violações da segurança em redes IP. A descrição não está vinculada a implementações ou produtos específicos, mas reúne as contramedidas em tipos ou categorias genéricos, oferecendo uma conceituação simples, uma breve descrição do seu funcionamento e uma rápida análise de suas possibilidades e limitações. As contramedidas apresentadas aqui foram selecionadas por seu potencial de eficiência ou por terem sido desenvolvidas a partir de conceitos cuja importância e validade tornaram-se fundamentais.

Os tipos de contramedidas descritos são as seguintes :

- **VPNs - Virtual Private Networks**
- **Firewalls**
- **Intrusion Detection Systems**
- **Protocolos seguros**

1. VPNs - Virtual Private Networks :

As VPNs – Virtual Private Networks são redes virtuais estabelecidas entre hosts remotos de modo que a comunicação entre eles sofra um estrito controle, resultando em garantias efetivas da autenticidade, confidencialidade e integridade dos dados transmitidos e recebidos.

- **Aplicabilidade :**

  A tecnologia das Virtual Private Networks tem encontrado grande aceitação no mercado e vem sendo adotada por empresas e instituições que encontram a necessidade, bastante usual aliás, de conectar sua rede local a redes remotas ou a hosts operando isoladamente fora dos limites físicos da rede local. Uma situação

típica em que a VPN pode ser de grande utilidade é representada pela empresa que mantém uma rede em sua matriz e necessita estabelecer conexão com as redes de uma ou mais filiais situadas em outras cidades. Outra circunstância típica de emprego de VPN é aquela em que um host isolado, como o notebook usado por um agente de vendas, precisa ter acesso à rede corporativa utilizando uma conexão discada a partir de localizações que variam dia a dia. Naturalmente deseja-se que esse acesso seja possível sem comprometimentos da segurança.

A comunicação por meio de uma VPN beneficia-se de uma série de funções implementadas pelos protocolos de comunicação sobre os quais a rede virtual é estabelecida. Alguns dos benefícios oferecidos pelas VPNs são :

- **Possibilidade de utilizar a Internet ou outra rede pública como um link de longa distância entre as redes ou hosts remotos. Essa possibilidade implica em redução de custos, uma vez que dispensa o estabelecimento de uma infra-estrutura de comunicação própria. Esta é, possivelmente, a justificativa mais comum para o emprego de VPNs.**

- **Possibilidade de comunicar dados de forma segura, com o apoio de recursos de criptografia e autenticação, por meio de uma ou várias redes supostamente inseguras.**

- **Possibilidade de comunicar dados através de várias redes heterogêneas. Esta possibilidade dispensa a adoção de um padrão único e confere grande flexibilidade aos projetos de expansão ou integração de redes.**

- **<u>Conceito de funcionamento</u> :**

O funcionamento de uma VPN baseia-se, em seu nível mais elementar, na possibilidade de inserir ou encapsular qualquer tipo de pacotes em outros pacotes, controlados por um protocolo que suporte funções de roteamento, criptografia e autenticação. Por serem roteáveis, esses pacotes podem ser transmitidos entre hosts localizados em redes distintas através de um ambiente de inter-redes, tipicamente

através da Internet, ainda que os pacotes encapsulados não sejam roteáveis. Por serem criptografados, os pacotes mantém-se seguros enquanto trafegam pelas redes inseguras. Por serem trocados através de uma conexão estabelecida mediante autenticação, fica assegurada a autenticidade dos dados a partir de uma origem conhecida.

- **<u>Descrição da Virtual Private Network</u> :**

As VPNs têm seu funcionamento vinculado ao uso de protocolos especiais. Dois dos protocolos mais utilizados na implementação de VPNs são o PPTP – Point-to-Point Tunneling Protocol e o L2FTP – Layer 2 Tunneling Protocol. Ambos têm sido objeto de um esforço deliberado de desenvolvedores e fabricantes de software no sentido de obter interoperabilidade a partir da convergência de suas especificações. Em função disso e em benefício da objetividade, focalizo a descrição da tecnologia de VPNs no protocolo PPTP, por considerar que essa abordagem, ainda que restritiva, é adequada à profundidade da análise conduzida aqui.

O PPTP é um protocolo de encapsulamento ponto a ponto, constituindo-se num padrão industrial aberto. O PPTP foi projetado com base no PPP - Point-to-Point Protocol - e no protocolo IP. O PPTP fornece recursos de autenticação e compressão de dados. O IP confere as capacidades de roteamento que permitem o tráfego orientado do PPTP pela Internet ou por redes similares.

O PPTP realiza o encapsulamento dos pacotes de uma sessão PPP em pacotes IP por meio de um protocolo de encapsulamento denominado GRE - Generic Routing Encapsulation. Adicionalmente, ele provê criptografia para os dados contidos nos pacotes.

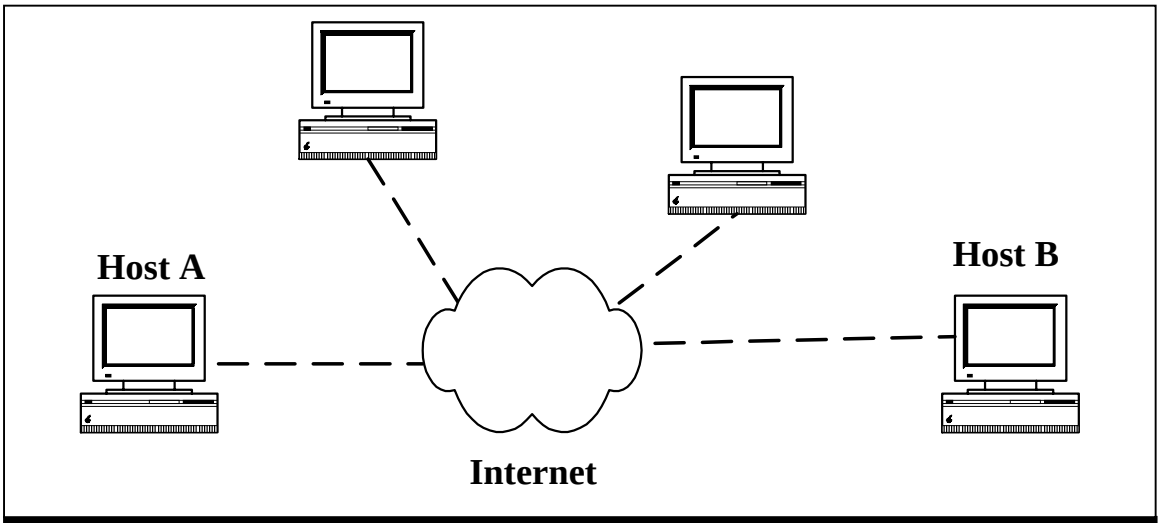
Como decorrência desse processo, a estrutura de um pacote PPTP pode ser representada como mostrado a seguir :

Cabeçalhos da camada Física	Cabeçalho IP	GRE	PPP	Dados .....

Nesta representação, devem ser entendido como Dados os pacotes encaminhados a partir da rede interna, incluindo-se seus cabeçalhos, sejam eles quais forem. Esse tratamento permite que protocolos diversos sejam encaminhados através de uma conexão PPTP, ainda que não apresentem características compatíveis com a rede intermediária por onde é estabelecida a conexão.

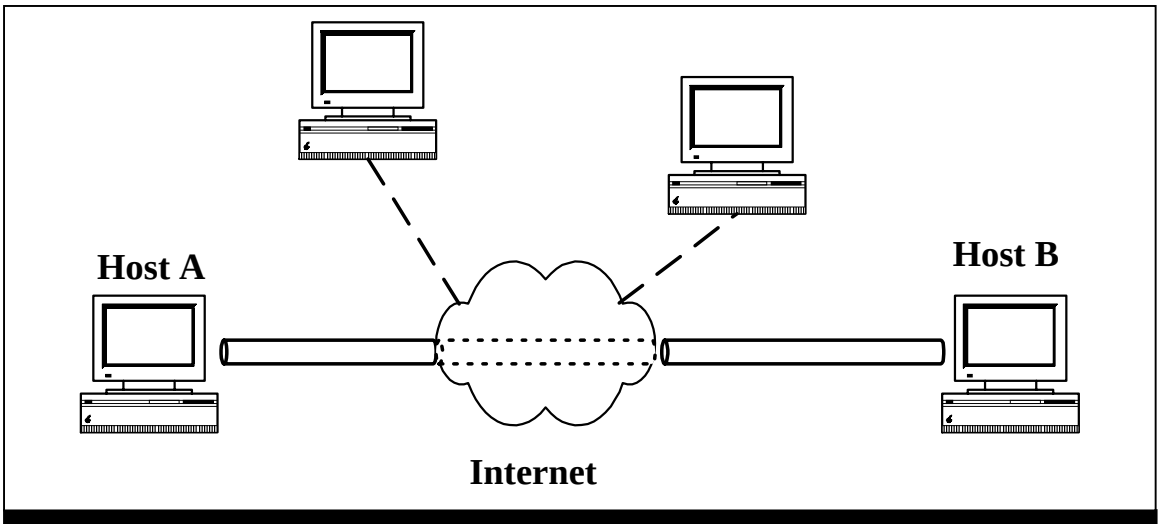
O PPTP utiliza dois tipos de mensagens : as mensagens de dados e as mensagens de controle. O tráfego dessas mensagens define dois canais separados, cada um deles com características próprias. O canal de dados baseia-se na utilização de um protocolo de transporte não confiável, como o UDP. Por ele são conduzidas as mensagens de dados, que contém os frames PPP encapsulados pelo PPTP. O canal de controle, por sua vez, é formado a partir de em uma conexão confiável gerenciada pelo próprio PPTP. Por esse canal trafegam mensagens de controle utilizadas para estabelecer, manter e finalizar a conexão lógica conhecida como túnel. O túnel é composto pelos canais de controle e de dados e pelas sessões estabelecidas através deles. A autenticação é uma das funções envolvidas no estabelecimento do túnel e é suportada por um protocolo específico como o MS-CHAP ou similar. A utilização de sequence numbers é outra característica das mensagens de controle que tem influência direta na segurança, restringindo as chances de montagem de um ataque de personificação bem sucedido contra a conexão que é gerenciada por seu intermédio.

Num cenário bastante simplificado, uma VPN pode ser estabelecida entre dois computadores conectados a uma rede compartilhada, como a Internet por exmplo. A rede é dita compartilhada porque outros hosts estão conectados a ela, além daqueles que irão comunicar-se por meio da VPN. A figura a seguir ilustra esse cenário :



VPN – Cenário anterior ao estabelecimento.

Em cada um dos hosts é habilitado o protocolo PPTP de modo a que os datagramas trocados entre eles sejam encapsulados e que um canal de controle garanta a conexão direta entre ambos. A rede virtual privada estabelecida funcionará como um túnel ligando os hosts através da Internet.

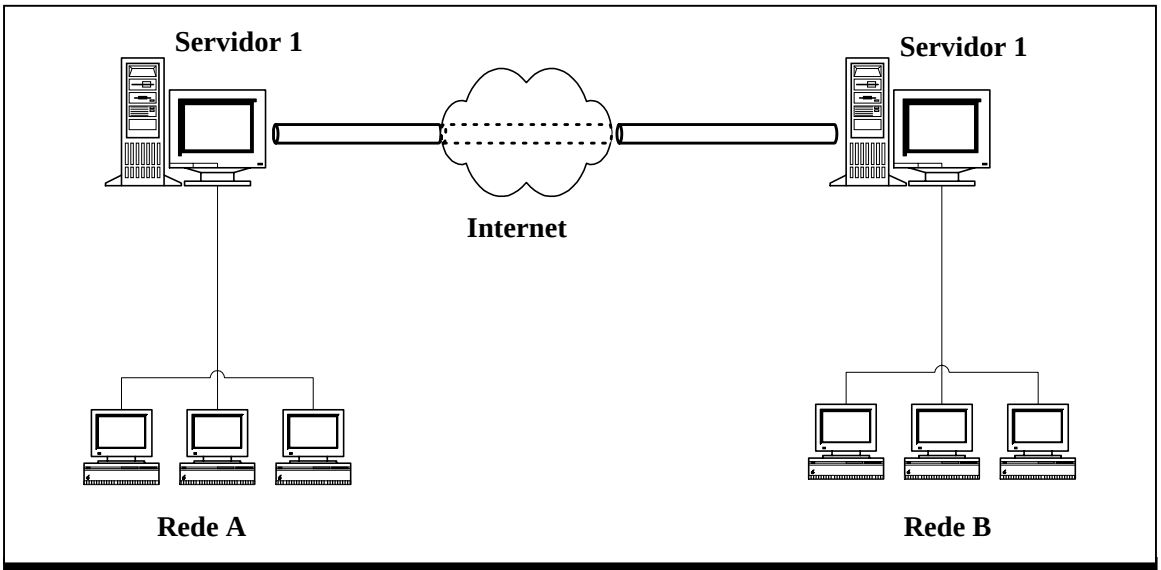


VPN – Estabelecimento do túnel através da rede compartilhada .

Expandindo essa idéia, considere-se que os dois hosts conectados pela VPN básica do exemplo anterior sejam servidores integrantes de redes locais diferentes. Esses servidores devem ter a capacidade de estabelecer uma conexão ponto a ponto entre si e reencaminhar pacotes recebidos de suas redes internas através dessa conexão.

Essa capacidade normalmente é fornecida pela implementação de serviços de roteamento.

A estrutura topológica resultante da implementação da VPN nessas condições é ilustrada na figura que se segue :



VPN estabelecida entre duas redes .

Nessa situação, os pacotes enviados por hosts da rede A para hosts de destino da rede B serão encapsulados pelo PPTP no Servidor 1 e transmitidos através da Internet até o Servidor 2, onde terão seus cabeçalhos IP externos descartados antes de serem encaminhados a seu destino na rede B.

Essa solução nos permite atender às necessidades da empresa fictícia, descritas no início deste tópico. A comunicação entre as redes A e B será suportada pela infra-estrutura pública. Os custos de implementação de infra-estrutura própria são suprimidos. Podem ser eliminados também custos relativos às tarifas de ligações interurbanas ou utilização de links de longa distância.

Ainda que as redes internas empreguem protocolos incompatíveis com o ambiente da Internet, como IPX ou NetBEUI, a comunicação será possível, já que os pacotes são encapsulados em datagramas IP. A forma como hosts e recursos estão

identificados tampouco irá requerer alteração, podendo ser mantidos os padrões habituais, normalmente associados aos protocolos em uso. O endereçamento dos protocolos internos também será oculto no encapsulamento.

O trânsito dos pacotes pela web não implica em uma vulnerabilidade, na medida em que a criptografia aplicada aos dados pode preservar sua confidencialidade. Parâmetros e informações de sistema contidos nos cabeçalhos dos protocolos internos podem ser também criptografados, negando aos eventuais atacantes a possibilidade de acesso a eles.

- **<u>Avaliação crítica</u> :**

As Virtual Private Networks impuseram-se como uma solução eficaz, capaz de atender a necessidades que freqüentemente são conflitantes entre si, como o estabelecimento de uma estrutura de comunicações segura e a limitação de custos.

Uma descrição sumária dessa tecnologia, como a que empreendo aqui, pode ratificar a noção comum de que a maior contribuição das VPNs para a manutenção da segurança baseia-se na criptografia de dados. Essa conclusão é correta apenas parcialmente. Além de preservar a confidencialidade e integridade dos dados pelo uso de recursos criptográficos, o emprego de VPNs é vantajoso por outras razões que merecem ser mencionadas.

Os procedimentos de autenticação envolvidos no estabelecimento do túnel constituem-se numa medida efetiva para evitar ataques de personificação e para garantir que os agentes envolvidos na comunicação são realmente quem afirmam ser.

O emprego de VPN oferece ao administrador ou designer de rede alternativas no sentido de implementar ou manter redes baseadas em protocolos diferentes de TCP/IP. A possibilidade de emprego da Internet como rede intermediária entre redes não IP evidentemente afasta, ainda que não elimine, todo um conjunto de ameaças à segurança especificamente relacionadas a esse protocolo. Em

determinadas situações, esse pode ser um fator de decisão a ser considerado no projeto de uma nova rede, ou de parte de uma rede que apresente requisitos especiais.

Num cenário em que o problema envolva a interligação de redes existentes baseadas em outros protocolos, poder manter o modelo correntemente em uso sem a substituição de protocolos e outras medidas associadas pode estender a validade de rotinas e métodos de manutenção da segurança. Essas condições contribuirão para reduzir a necessidade de reconfigurações e os custos de administração das redes, tendo um impacto apreciável, ainda que indireto, sobre a segurança.

As limitações e vulnerabilidades apresentadas pelas VPNs estão associadas, em sua maioria, a deficiências dos recursos criptográficos utilizados nas implementações dos protocolos de encapsulamento. A divulgação de fraquezas nos métodos de criptografia implementados com o PPTP da Microsoft teve grande repercussão e exemplifica bem essa classe de falhas. A descrição dessas brechas pode ser obtida em http://www.counterpane.com/pptp-faq.html.

Outras falhas, entretanto, podem ser observadas. A vulnerabilidade do canal de controle a ataques de negação de serviços é uma das mais relevantes, não só por seus efeitos diretos, mas também pelo potencial de emprego como etapa intermediária em um ataque mais complexo.

Segundo algumas estimativas, as vulnerabilidades do canal de controle devem suceder as fraquezas de criptografia como foco de interesse no desenvolvimento de técnicas de ataque orientadas à violação da segurança em VPNs. As possibilidades, ainda bastante teóricas, de obtenção de informações críticas e de personificação de hosts a partir do monitoramento e interferência nas comunicações desse canal certamente serão exploradas no futuro.

2. Firewalls :

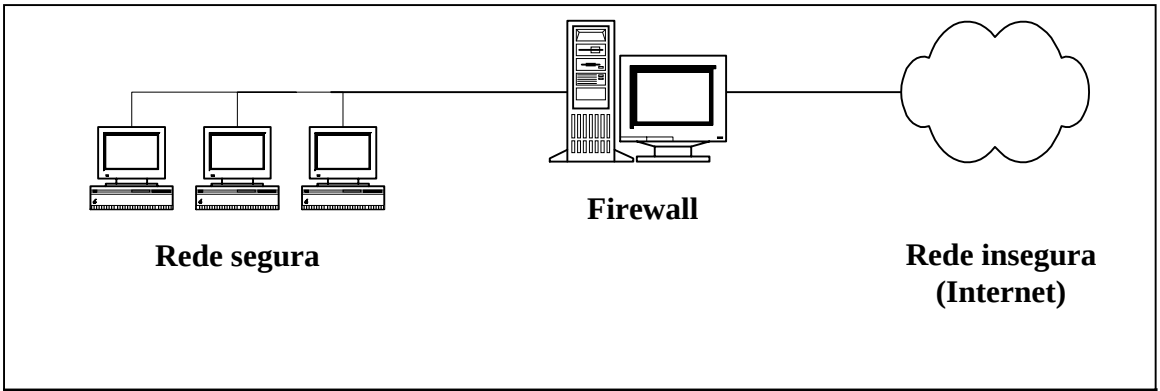
Um firewall é uma combinação de componentes de hardware e software utilizado para controlar tráfego e acesso entre duas redes, uma delas de interesse privado e considerada segura e outra pública e supostamente insegura. Como exemplo de rede insegura podemos tomar a Internet ou um setor da rede corporativa onde os requisitos de segurança são mais baixos.

- **Aplicabilidade :**

  Um cenário típico de emprego proveitoso de um firewall é representado pela situação de uma empresa cujas atividades envolvem tanto a disponibilização de dados e serviços a usuários externos através da Internet quanto o fornecimento de acesso controlado à Internet e a outras redes remotas para os usuários de sua rede interna.

- **Conceito de funcionamento :**

  A modalidade mais simples de firewall seria constituída de um computador dotado de duas interfaces de rede, uma delas conectada à rede interna ou segura e outra conectada à rede externa ou insegura. Habilitando-se no computador um conjunto de funções de roteamento e mecanismos de controle do tráfego seria possível encaminhar pacotes nos dois sentidos entre as redes interna e externa e administrar o tráfego que passa por este firewall elementar, filtrando-o ou bloqueando-o conforme regras predefinidas.



Firewall composto por um computador com duas interfaces de rede- .

É claro que na prática a implementação e operação de firewalls pode envolver uma complexidade muito maior do que a apresentada por este modelo conceitual. Conforme o caso, um firewall pode ser constituído por uma combinação de diversos equipamentos e componentes de software, reunidos num único componente de hardware ou organizados numa estrutura topológica adequada ao desempenho de suas funções.

- **<u>Descrição de Firewalls</u> :**

Existem componentes de firewall típicos que se distinguem por suas funções e por seu método de funcionamento. A seguir discutiremos brevemente os roteadores de filtragem de pacotes, os gateways de nível de aplicativo, também conhecidos como proxies, e os gateways de nível de circuito:

a. Roteador de filtragem de pacotes :

Uma implementação simples de firewall é construída sobre um roteador capaz de examinar os parâmetros presentes nos cabeçalhos dos pacotes que recebe e, baseado em regras preestabelecidas, decidir pelo encaminhamento desse pacote à rede de destino, interna ou externa, ou pela eliminação do pacote.

As regras observadas por um roteador de filtragem de pacotes baseiam-se em informações que podem ser obtidas a partir dos cabeçalhos dos pacotes recebidos.

Essas informações normalmente incluem :

- **Endereço IP de origem**
- **Endereço IP de destino**
- **Porta TCP ou UDP de origem**
- **Porta TCP ou UDP de destino**
- **Tipo de mensagem ICMP**
- **Informações sobre o protocolo de camada superior**

Baseado nesse conjunto de informações é possível definir regras de filtragem, que poderiam ser enquadradas em três categorias genéricas :

- **Filtragem baseada em origem e destino - a informação sobre os endereços IP de origem e destino fornece indicações sobre os hosts que estão envolvidos na comunicação. A partir dessa informação, é possível estabelecer regras que autorizem ou neguem o acesso a determinados hosts, conforme as necessidades de segurança.**

- **Filtragem baseada em serviço – a informação sobre as portas de origem e destino fornece indicações sobre qual serviço está sendo conectado. A partir dessa informação, pode ser implementada uma regra que negue a utilização de determinados serviços considerados inseguros ou indesejáveis.**

- **Filtragem avançada – as regras de filtragem avançada podem combinar informações referentes a origem, destino e serviços ou ainda incluir verificações especiais como integridade de datagramas IP ou numeração de fragmentos.**

Há um número considerável de limitações para a técnica de filtragem de pacotes. Por exemplo, os serviços que usam portas diferentes em cada conexão, como é o caso do NFS e de serviços baseados em UDP, não podem ser filtrados convenientemente.

Outra limitação diz respeito à dificuldade em definir regras de filtragem baseadas em parâmetros relativos às camadas superiores, como a camada de aplicação. As informações dessas camadas que podem ser observadas na filtragem de pacotes são reduzidas e isso restringe a funcionalidade desse método. Há uma possibilidade que exemplifica claramente essa limitação. Um atacante posicionado em uma rede externa pode pretender inserir numa máquina da rede interna um applet ou outro programa que possa ser executado mais tarde, causando danos ao sistema alvo. Isso é possível porque o roteador de

filtragem de pacotes não será capaz de analisar o conteúdo dos dados recebidos e identificar o componente nocivo que está sendo enviado à rede interna.

b. Gateway de nível de aplicativo ou proxy :

Um proxy oferece uma proteção bem mais sofisticada e eficaz que um roteador de filtragem de pacotes. O conteúdo referente à camada de aplicação pode ser observado e interpretado, permitindo a implementação de regras de bloqueio e acesso mais completas.

Um proxy é normalmente um serviço executado em um servidor posicionado entre as redes interna e externa. Cada conexão solicitada entre hosts dessas duas redes é estabelecida como duas conexões separadas, de cada um dos hosts com o proxy.

O roteamento de pacotes IP é normalmente desabilitado no proxy, que não reencaminha pacotes entre as duas redes, mas gera pacotes de acordo com as necessidades de cada conexão. Os hosts interno e externo não chegam sequer a conhecer os endereços IP um do outro. Para cada um deles a conexão é estabelecida com o endereço IP da interface de rede correspondente no servidor proxy. Isso é muito conveniente porque permite ocultar endereços IP na rede interna da observação externa.

Como cada host deve conectar-se a uma porta no proxy para comunicar-se com o host da outra rede, é possível implementar mecanismos de autenticação que acrescentam um nível adicional de segurança.

Os proxies representam uma solução eficiente e flexível e, relativamente, apresentam poucas limitações. Uma delas consiste num eventual atraso na comunicação, que passa a depender de um processo intermediário mais complexo e demorado. Esse atraso pode gerar falhas decorrentes de time out ou dificultar o uso de aplicações específicas, sobretudo se o servidor sobre o qual o proxy é executado não estiver em condições de lidar com a sobrecarga

produzida pelo serviço. Outra limitação é representada pela necessidade de reconfiguração de clientes para acessar a rede externa via proxy. Além de significar um encargo administrativo adicional, em determinadas topologias existe a possibilidade de que alguns clientes sejam reconfigurados para acessar a rede externa diretamente, com o comprometimento da segurança.

c. Gateway de nível de circuito :

Gateways de nível de circuito funcionam como proxies simplificados, intermediando conexões entre as redes interna e externa com funções de controle reduzidas. Essas funções geralmente limitam-se à autenticação do usuário.

A exemplo dos proxies, os gateways de nível de circuito normalmente são implementados como serviços em execução em um servidor de rede. Um exemplo clássico é o serviço SOCKS, amplamente utilizado em servidores UNIX.

A par das limitações inerentes à simplicidade de seus métodos de funcionamento, as gateways de nível de circuito oferecem grande flexibilidade, tendo seu uso indicado em situações nas quais os requisitos de controle são reduzidos ou quando é necessário o suporte a serviços baseados em UDP.

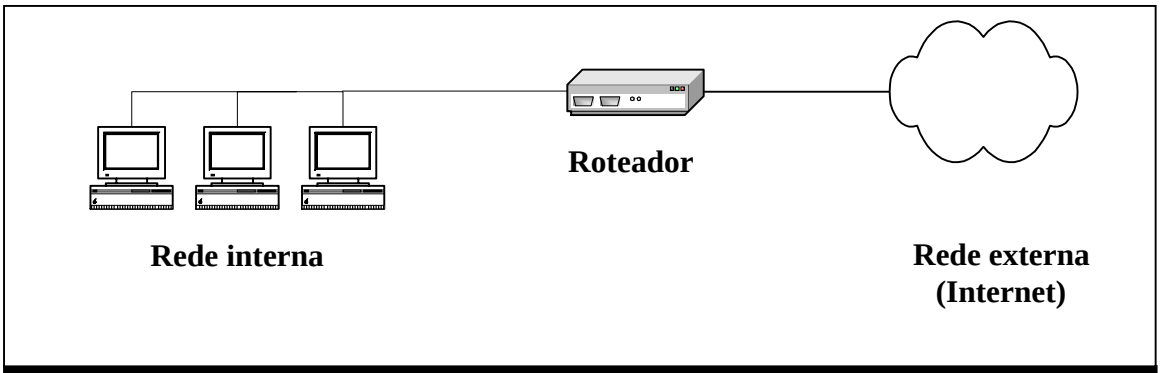
O planejamento de um firewall deve procurar combinar os componentes descritos acima numa estrutura adequada às necessidades de controle e segurança da rede. Um problema comum consiste em estabelecer condições de restrição de acesso eficazes sem que os limites impostos para o acesso a recursos da rede interna ou da Internet sejam superdimensionados.

Embora os componentes de um firewall possam ser arranjados de formas bastante diversas, há quatro modelos genéricos cujas características iremos descrever. São eles :

a. **Firewall de filtragem de pacotes**
b. **Firewall de gateway dual-homed**
c. **Firewall de host encoberto**
d. **Firewall de sub-rede encoberta**

a. Firewall de filtragem de pacotes :

É o mais simples dos modelos apresentados aqui. Consiste apenas em um roteador de filtragem de pacotes posicionado entre as redes interna e externa, como mostra o esquema abaixo :

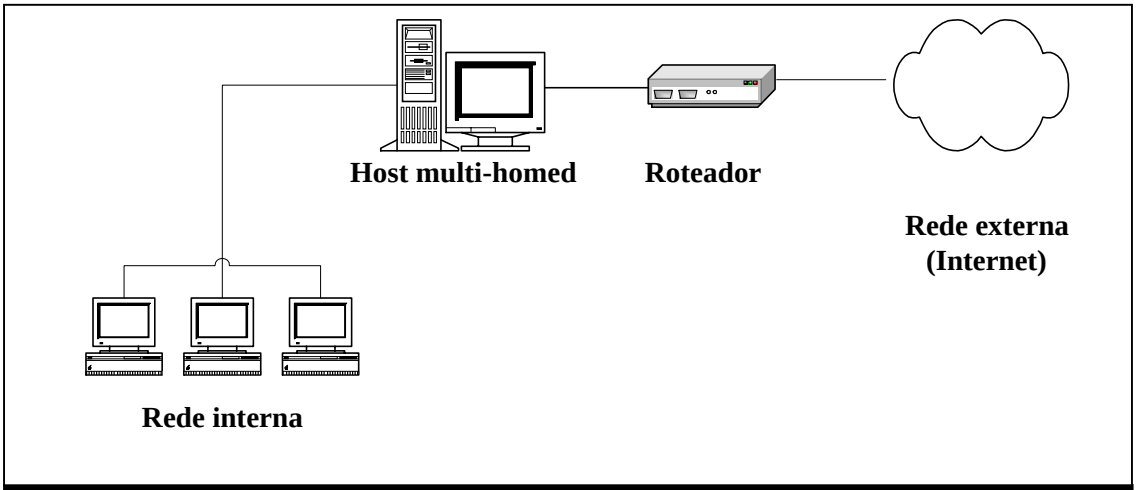


Firewall de filtragem de pacotes .

Este modelo é também o que provê o menor nível de segurança e as funcionalidades mais limitadas. Um firewall de filtragem de pacotes possibilita apenas o controle de acesso baseado em origem e destino dos pacotes e nos serviços acessados. Funções de controle adicionais, como autenticação de usuários, devem ser implementadas em cada host da rede interna.

b. Firewall de gateway multi-homed :

Este modelo é composto de um roteador e de um host multi-homed dispostos conforme ilustrado abaixo :



Firewall de filtragem de pacotes .

O roteador não executa qualquer filtragem de pacotes. Todo o tráfego recebido da rede externa é reencaminhado sem restrições. O host multi-homed é um host dotado de duas ou mais interfaces de rede. A cada interface de rede é associado um endereço IP diferente. Os serviços que fornecem roteamento de datagramas IP são desabilitados no host, bloqueando o tráfego entre as redes interconectadas, que passa a ser condicionado pela aplicação das regras estabelecidas por um serviço de proxy ou de gateway em nível de circuito.

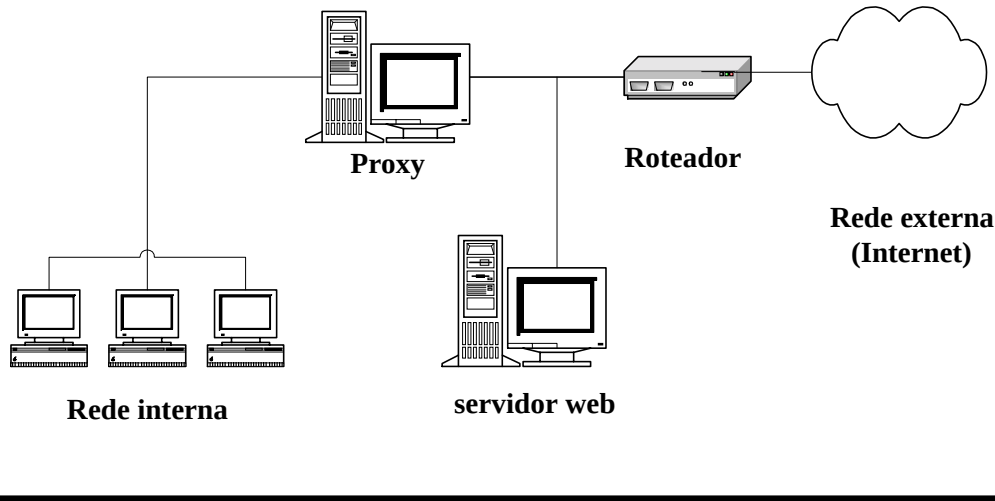
O modelo define uma área da rede local, situada entre o host multi-homed e o roteador, na qual as restrições ao acesso externo não são aplicáveis. Nessa área podem ser posicionados hosts que oferecem serviços com baixos requisitos de segurança e que devam receber requisições frequentes de fora da rede local, como um servidor web, por exemplo. Tal arranjo liberaria os serviços de proxy

executados no host multi-homed da sobrecarga associada ao controle do acesso externo a esses hosts.

O nível de segurança oferecido por esse modelo deve ser satisfatório para a maioria das situações. Ele oferece ainda a flexibilidade adicional de posicionar hosts em uma área de acesso facilitado.

c. Firewall de host encoberto :

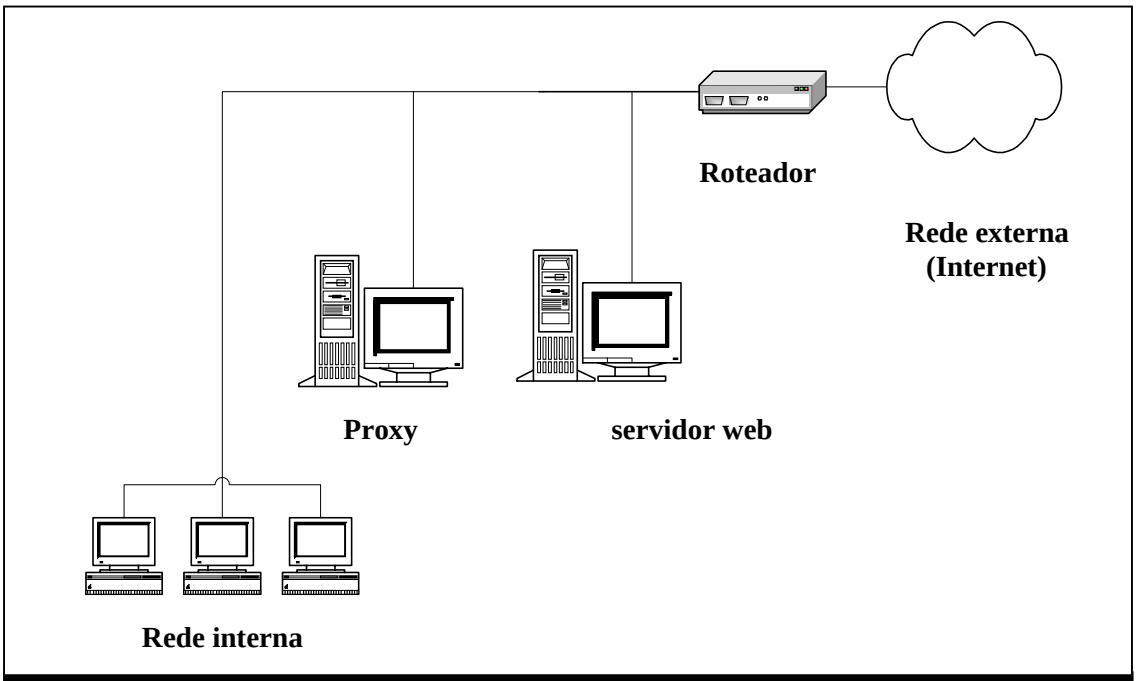
Este modelo reúne um roteador de filtragem de pacotes e um host multi-homed, usualmente denominado servidor proxy ou ainda bastion host, no qual são executados serviços de proxy ou de gateway de nível de circuitos. A figura abaixo exemplifica o modelo :



Firewall de host encoberto (1) .

O roteador é configurado para encaminhar os pacotes originados da rede externa para o servidor proxy ou, conforme o caso, para hosts situados na área intermediária da rede, como um servidor web destinado a ser acessado por usuários externos, por exemplo.

O modelo admite uma variação em que o host multi-homed é substituído por um host de uma só interface de rede, conforme ilustrado na figura a seguir :



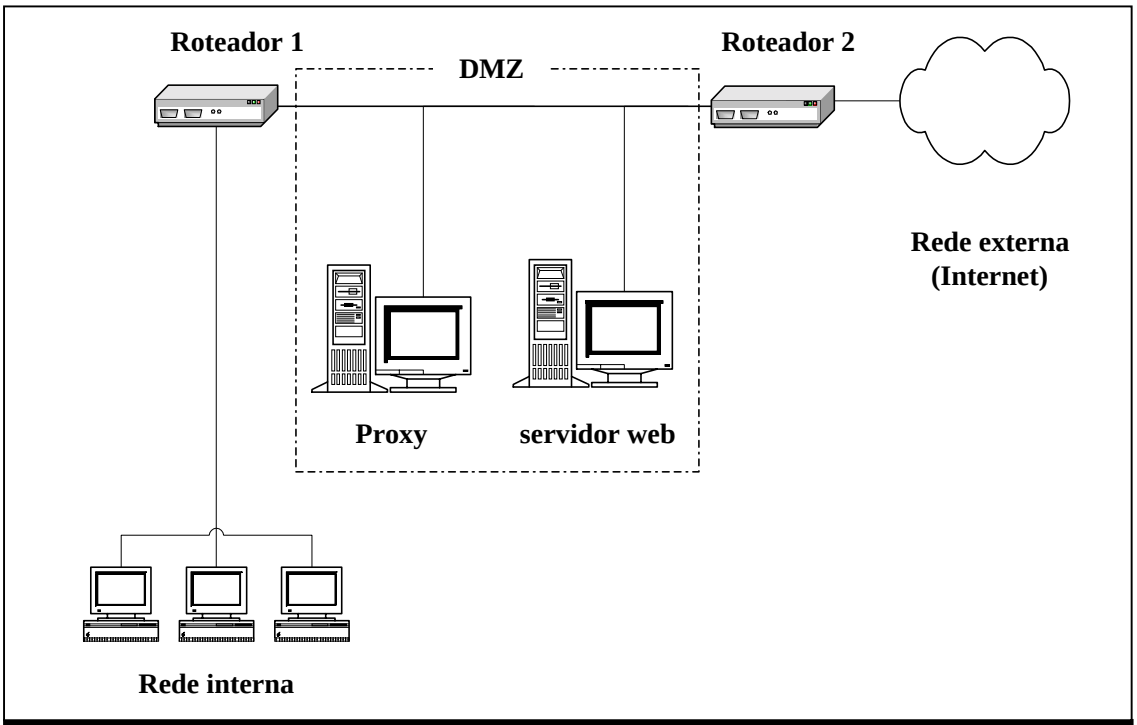
Firewall de host encoberto (2) .

Neste caso o acesso de usuários internos à rede externa pode ser feito diretamente através do roteador ou usando os serviços do servidor proxy. As configurações de cada cliente e as regras de filtragem implementadas no roteador irão combinar-se para determinar, em cada caso, qual alternativa será efetivada.

O modelo oferece a vantagem de compor as funcionalidades da filtragem de pacotes com as do serviço de proxy, provendo um nível superior de segurança. Mantendo o servidor proxy encoberto pelo roteador, sua exposição a investigações e ataques é reduzida, o que contribui também para tornar mais robusta a solução. Em sua segunda variação, o modelo apresenta ainda a flexibilidade de permitir formas diferenciadas de acesso à rede externa para os clientes da rede interna, o que possibilita maior rapidez de fluxo e balanceamento da carga imposta aos serviços de proxy.

d. Firewall de sub-rede encoberta:

Este modelo emprega um proxy situado entre dois roteadores de filtragem de pacotes, conforme representado na ilustração a seguir :



Firewall de sub-rede encoberta .

O modelo define uma área entre os roteadores que é usualmente chamada de DMZ – Demilitarized Zone. Nessa área é estabelecido um nível de segurança intermediário, dependente apenas da filtragem de pacotes realizada pelo roteador externo. Isso torna a DMZ uma localização adequada para recursos com previsão de acesso externo frequente, como servidores de acesso remoto e servidores web.

O firewall de sub-rede encoberta provê o mais elevado nível de segurança dentre os modelos descritos aqui. O estabelecimento de regras de filtragem nos roteadores que canalizem para o servidor proxy tanto o tráfego proveniente da rede interna quanto o que se origina na rede interna empresta grande consistência ao modelo. A existência da DMZ como uma área bem definida impõe uma dificuldade considerável para obter acesso não autorizado entre as redes interna e externa. Apesar disso, o modelo é flexível o bastante para permitir que hosts específicos sejam acessados sem a intermediação do servidor proxy, possibilitando economia de recursos e comunicação mais rápida.

- **<u>Avaliação crítica</u> :**

Além das limitações e possibilidades específicas de cada modelo genérico de firewall, mencionadas no âmbito das descrições empreendidas anteriormente, existem outras capacidades e fatos concernentes ao funcionamento e emprego dos firewalls que devem ser mencionados.

Um deles diz respeito à opção entre firewalls integrados e firewalls montados a partir de componentes autônomos. Segundo minha análise, os firewalls integrados, que em alguns casos são representados por hardware e software dedicados reunidos em uma única caixa, podem ser capazes de fornecer um conjunto mais consistente de recursos e métodos, o que deve implicar em um nível superior de segurança. Em contrapartida, o emprego isolado desse tipo de solução conduz normalmente ao estabelecimento de um ponto único de falha, deixando a segurança da rede, e em certas circunstâncias a sua própria funcionalidade, na dependência da solidez e estabilidade da solução adotada. Os firewalls compostos por elementos agregados como roteadores e servidores representam uma alternativa que pode oferecer maior resiliência e flexibilidade, apesar de apresentar custos potencialmente mais altos.

Outro aspecto a ser considerado refere-se à implementação de firewalls baseados em serviços executados sobre plataformas comerciais não dedicadas, o que alguns autores chamaram de OS Shields. A redução de custos tem sido uma justificativa habitual para essa escolha. O Microsoft Proxy Server sobre Windows NT e o ISA (Internet Security and Acceleration) Server sobre Windows 2000 exemplificam perfeitamente a situação a que me refiro. Nesses casos as vulnerabilidades do sistema operacional representam um risco potencial à segurança do firewall e, por extensão, de toda a rede. Essas vulnerabilidades podem ser ampliadas se, em nome ainda da redução de custos, optar-se por executar outros serviços num host que detenha a responsabilidade por serviços associados ao funcionamento do firewall.

Mais uma característica dos firewalls que precisa ser analisada em uma avaliação é a disponibilidade de funções complementares. O registro de eventos como eliminação de pacotes e tentativas de acesso não autorizado é em geral de grande utilidade para

a detecção de um ataque em preparação ou andamento. A capacidade de gerar relatórios claros e consistentes a partir desses registros complementa sua utilidade. Ferramentas de auditoria integradas ao software do firewall são um recurso altamente desejável, que valoriza qualquer solução ou produto.

Podemos obsevar ainda se o firewall detém a capacidade de reagir a determinados eventos, oferecendo resposta a uma ameaça em tempo hábil de neutralizá-la sem que sua funcionalidade ou o conjunto de funcionalidades da rede sejam afetados. Tal capacidade pode depender de funções avançadas ou da redundância de recursos e pode ser obtida como parte integrante de determinadas soluções ou ser implementada a partir da adoção de uma arquitetura específica para o sistema que irá constituir o firewall.

O primeiro caso pode ser exemplificado pelas configurações de alarme disponíveis em diversas soluções de firewall. Essas configurações permitem definir limites para eventos específicos em termos de intervalo de tempo ou número de ocorrências. Se um evento como uma varredura de portas for observado pelo firewall uma única vez, não haverá reação desnecessária que consuma recursos ou forneça ao atacante indícios sobre a existência e atividade do firewall. Porém, se o evento repetir-se várias vezes e o limite configurado for atingido, sua ocorrência será notificada pelos mecanismos de alarme existentes.

O segundo caso é ilustrado pela utilização de um array de proxies formado por vários servidores associados entre si. Caso um desses servidores seja atingido por um ataque de negação de serviço e venha a tornar-se inoperante, os demais servidores integrantes do array serão capazes de redistribuir entre si as solicitações de conexão e outras tarefas, mantendo a disponibilidade do conjunto.

As soluções de firewall evoluem no sentido de tornarem-se sistemas mais completos, incorporando capacidades que lhes permitam reagir automaticamente a falhas e ameaças. Dispensando a intervenção direta de operadores, o tempo de resposta a eventos dessa natureza é reduzido, contribuindo para tornar ineficazes muitas técnicas de ataque e para assegurar aos serviços da rede um alto grau de

disponibilidade e confiança. Essa evolução aproxima os firewalls de um conceito mais amplo de sistema ao qual pode ser conferida a responsabilidade de detectar, analisar e impedir invasões e ataques à rede, de forma rápida e automatizada . Essa classe de sistemas, que mereceu a designação de IDS – Intrusion Detection System, já conta com representantes em uso efetivo. Ela será analisada no próximo tópico deste trabalho.

O emprego de um firewall pode ser suficiente para garantir a segurança do perímetro da rede, mas certamente não será capaz de impedir que ameaças originadas dentro do perímetro alcancem efetividade. Ameaças internas são historicamente mais frequentes do que ataques lançados de fora para dentro e costumam encontrar condições favoráveis para atingir não só hosts comuns na rede local, mas também firewalls cujas configurações refletem uma preocupação exclusiva com as ameaças exteriores.

Finalmente, deve ser lembrado que um firewall poderoso e bem configurado não é mais do que um dos componentes do sistema de defesa da rede. É indispensável associar a utilização do firewall a outras contramedidas de caráter técnico e administrativo, a fim de estabelecer condições sólidas de segurança.

3. <u>Intrusion Detection Syatems</u> :

"A detecção de invasão é melhor compreendida como uma capacidade, não como uma única ferramenta. "

Stephen Northcutt

Os IDS – Intrusion Detection System surgiram da necessidade de uma solução integrada capaz de monitorar e avaliar a diversidade de eventos que podem estar relacionados ao comprometimento da segurança de uma rede. Iniciativas pioneiras de desenvolvimento e utilização de IDS foram tomadas pela Marinha dos Estados Unidos e outros órgãos governamentais daquele país. Cada uma dessas iniciativas levou a resultados que

incorporavam métodos e recursos próprios. Assim, a possibilidade de padronização dessa tecnologia foi comprometida logo nos estágios iniciais de seu desenvolvimento.

O interesse do setor governamental americano em soluções desse tipo de pronto correspondeu, como em outras ocasiões, a um esforço de centros de pesquisa no sentido de aperfeiçoá-las. Logo chegaram ao mercado soluções comerciais que colocavam à disposição de empresas e instituições privadas recursos de detecção e análise adequados aos ambientes de rede padronizados.

Atualmente os IDS são uma das soluções mais valorizadas no mercado, porque suportam o conceito, já amplamente aceito, de integração de medidas na manutenção da segurança.

- **<u>Aplicabilidade</u> :**

  A variedade de soluções IDS disponíveis atualmente corresponde um espectro de utilização que transcende o típico ambiente institucional ou corporativo onde essa tecnologia originou-se. Na verdade, os IDS têm aplicação útil e vantajosa em praticamente todos os ambientes em que exista uma rede conectada a outras redes. O porte da rede, seus requisitos de segurança, seus perfis de utilização e as tendências de expansão futura são características que devem ser levadas em conta na seleção do sistema adequado.

- **<u>Conceito de funcionamento</u> :**

  Um IDS consiste, em termos conceituais, num sistema integrado capaz de detectar tentativas de comprometimento ou de acesso indevido a uma rede ou seus recursos a partir da execução de três tipos de funções :

  - **Monitoramento e registro de eventos**
  - **Análise dos eventos registrados**
  - **Execução de medidas ou procedimentos de resposta automática**

Essas funções genéricas guardam entre si relações de dependência e sequencialidade. Medidas de resposta automática precisam ser desencadeadas e orientadas a partir da análise dos eventos registrados. Esta, por sua vez, será tão eficiente quanto tiver sido o monitoramento dos eventos.

É verdade que as soluções mais simples limitam-se a executar uma ou duas dessas funções. Podemos começar a avaliar a capacidade de um IDS e sua adequação a um problema de segurança real questionando quais das funções acima ele é capaz de executar.

- **<u>Descrição de Intrusion Detection Systems</u> :**

Pelo que já dissemos a respeito da variedade de características que pode ser observada nos IDS disponíveis atualmente, é difícil elaborar uma descrição desses sistemas que seja perfeitamente aplicável a cada um deles, sem incorrer em exageros ou omissões. Por isso, expomos aqui um conjunto de métodos e funcionalidades que tornam claro como funcionam os IDS e podem ser encontrados nas maioria deles, ainda que em graus muito variados de aperfeiçoamento.

Um IDS executa o monitoramento de eventos a partir de sensores ou agentes. Estes são componentes de software instalados em computadores e equipamentos de conexão. Normalmente, um agente é capaz de observar a ocorrência de eventos específicos, como solicitações de conexão e recebimento de datagramas, registrando informações de interesse sobre cada um desses eventos. São informações de interesse usualmente coletadas os endereços de origem de solicitações de conexão e certos parâmetros dos cabeçalhos IP dos datagramas recebidos. Os agentes costumam armazenar as informações coletadas por um período de tempo determinado. Ao final desse período, ou quando forem solicitados a fazê-lo, os agentes enviam as informações coletadas a uma entidade central de gerenciamento. É comum o uso de canais seguros para o envio dessas informações.

Na entidade central de gerenciamento ou console de gerenciamento as informações coletadas pelos agentes são registradas em uma base de dados para efeito de

comparação e análise. O formato dessa base de dados e os mecanismos utilizados em seu gerenciamento na console terão impacto decisivo na capacidade do IDS para processar as informações recebidas, associando eventos correlacionados e identificando tentativas de ataque e comprometimentos da segurança. A console deve ser capaz de gerar relatórios ordenados dos eventos, facilitando sua interpretação pelos analistas de segurança. A eficiência da console em produzir tais relatórios irá contribuir para a elaboração de análises rápidas e precisas.

Por fim, nos casos em que a análise dos dados indicar uma quebra da segurança consumada ou em andamento, podem ser emitidas instruções a agentes específicos para que sejam tomadas providências destinadas a neutralizar a ameaça ou a evitar seu agravamento. Dentre essas providências podem figurar desde uma nova coleta de informações em bases mais detalhadas até o bloqueio das solicitações originadas de um ou mais endereços IP ou a completa desativação de serviços supostamente comprometidos. Em sistemas mais sofisticados, esses procedimentos de resposta podem ser desencadeados de forma automática, baseado em critérios de avaliação preestabelecidos. Isso dispensa a intervenção direta de um analista, iniciando medidas reativas num tempo mais curto.

O funcionamento dos IDS, aqui descrito de forma esquemática, admite muitas variações. Essas variações podem ser melhor compreendidas a partir de uma classificação simples, que discrimina os IDS com base nos pontos ou setores da rede que ele observa e ainda nos métodos utilizados para identificar eventos de interesse. Assim, no que diz respeito à forma de monitoramento, teremos os seguintes tipos de IDS :

- **host based – um IDS deste tipo tem agentes instalados em hosts específicos e observa os eventos que envolvem esse host diretamente. Ele não considera o fluxo de dados na rede, concentrando-se exclusivamente no nó sob sua responsabilidade.**

- **network based – por oposição ao tipo anterior, os sistemas network based monitoram o tráfego da rede, registrando eventos relacionados à rede como**

um todo, sem concentrar-se em hosts determinados. Um sistema deste tipo em geral exige diversos agentes distribuídos pela rede. Esses agentes devem reportar inclusive a ocorrência de eventos que, se tomados isoladamente, podem não indicar claramente uma falha de segurança. Solicitações de conexão a determinados hosts ou a portas específicas são um bom exemplo. Caberá à console combinar as informações enviadas pelos agentes e verificar se elas revelam uma ameaça efetiva.

Já quanto ao método de identificação de eventos, podemos classificar os IDS da seguinte forma :

- **signature based – este tipo de IDS identifica os eventos de interesse por meio de características bem conhecidas das técnicas de ataque, as chamadas assinaturas. A assinaturas constam de uma base que deve ser atualizada periodicamente. Sistemas comerciais costumam dispor de atualizações distribuídas por seus fabricantes.**

- **traffic based – os IDS baseados em tráfego, por sua vez, identificam eventos a partir de alterações e anomalias no tráfego da rede.**

Em qualquer desses casos, há um certo número de fatores que condicionam o funcionamento do sistema. A seguir analisamos alguns deles :

- **quantidade de agentes – é fácil entender que um sistema, sobretudo um sistema baseado em rede, que disponha de um número maior de agentes estará em condições de observar mais eventos, aumentando com isso a possibilidade de detectar uma intrusão. A redundância de meios oferecida por múltilos agentes é também uma característica desejável num IDS. Deve ser considerado entretanto que um número excessivo de agentes pode produzir uma quantidade de informações além da capacidade de análise do sistema, prejudicando seu funcionamento.**

- **posicionamento dos agentes – a capacidade dos agentes em obsevar eventos é afetada diretamente por seu posicionamento na rede. Espera-se que a maioria das tentativas de intrusão incida inicialmente sobre roteadores e hosts situados no perímetro da rede. A DMZ é também uma área onde os eventos de interesse serão observados com mais frequência.**

- **capacidade de crítica – nem todos os eventos observados pelos agentes podem ser considerados anomalias ou indícios de invasão. Muitos deles, como solicitações de conexão e recebimento de mensagens ICMP, ocorrem regularmente como parte do funcionamento normal da rede. Quando o sistema interpreta um evento normal como uma violação da segurança temos um falso positivo. Se a cada falso positivo for acionado um alarme ou procedimento de resposta, o sistema pode ser levado a condições de funcionamento insustentáveis. O IDS deve ser capaz de distinguir entre a ocorrência normal dos eventos e as circunstâncias em que eles indicam uma invasão ou um ataque de negação de serviços. Essa crítica pode ser feita com base não só na natureza dos eventos, mas também pela sua origem, pela frequência com que ocorrem e por uma série de outras características. Como regra geral, quanto mais sofisticada for a capacidade de crítica do sistema, mais eficiente ele será.**

- **capacidade de convergência – como uma extensão da capacidade de crítica, a capacidade de convergir informações recebidas de vários agentes e de analisá-las sob uma perspectiva de conjunto possibilta ao IDS detectar uma variedade maior de tentativas de intrusão e com maior antecedência.**

- **recursos de resposta automática – em muitos sistemas estão disponíveis recursos que permitem acionar automaticamente certos procedimentos de defesa quando uma ameaça à segurança for identificada. O acionamento desses procedimentos é executado após uma análise positiva das informações, indicando um efetivo comprometimento. Normalmente é feita também uma comparação das ocorrências verificadas com limites de tolerância pré-configurados, a fim de evitar os efeitos de falsos positivos.**

  Entre os procedimentos de defesa que podem ser acionados pelo IDS estão :

- **interrupção temporária de conexões, por meio de mensagens RST.**

- **interrupção permanente de conexões determinadas, mediante o bloqueio de endereços de origem suspeitos.**

- **interrupção permanente de conexões, mediante a desativação de hosts ou serviços.**

- **reconfiguração de rotas, com o propósito de desviar o ataque para honeypots ou para sistemas alternativos.**

- **ativação de dispositivos especiais como sniffers ou ferramentas de análise.**

Os recursos de resposta automática não são capazes de oferecer um grau elevado de precisão e não substituem em absoluto a análise cuidadosa que pode ser executada por analistas de segurança devidamente qualificados e familiarizados com o ambiente. Contudo eles contribuem para a manutenção da segurança, possibilitando uma vigilância proativa e contínua e reduzindo o tempo de reação a ameaças.

- **<u>Avaliação crítica</u> :**

Cada tipo de IDS revela uma abordagem específica, à qual correspondem vantagens e desvantagens. Comparando a atuação dos sistemas host based e network based somos levados a concluir que os primeiros têm um foco mais definido e a capacidade potencial de gerar e analisar informações com maior rapidez e detalhamento. São por isso mais adequados ao emprego em nós de missão crítica, como servidores e roteadores. Já os sistemas network based fornecem um controle mais amplo, com abrangência sobre toda a rede, ficando em condições de detectar ameaças que incidem sobre vários hosts de forma sequencial ou simultânea. Isso os coloca em condição de comparar dados e estabelecer relações para detectar tentativas de intrusão em seus estágios iniciais, quando normalmente o atacante executa uma

sondagem em vários pontos da rede, buscando um alvo adequado. A essa capacidade contrapõe-se a desvantagem de gerar um volume maior de dados que em redes grandes pode tornar-se difícil de analisar.

Da comparação entre os sistemas baseados em tráfego com os que utilizam assinaturas é possível notar que os primeiros estão em condições de detectar, ou pelo menos de fornecer indícios de uma gama mais ampla de sondagens, invasões e comprometimentos, oferecendo ainda um grau de antecipação mais elevado. Exigem porém um esforço maior de análise de dados, o que pode impor uma demora indesejada à obtenção de resultados.

Já os sistemas signature based são capazes de resultados mais rápidos e precisos, em decorrência de sua forma de atuação mais direta. Sofrem contudo da dependência de uma base de assinaturas abrangente e atualizada, numa situação análoga à que experimentam os programas anti-vírus.

Uma forma de compensar as limitações de cada tipo de sistema consiste no emprego combinado de ferramentas de vários tipos. Essa alternativa só alcançará uma eficácia significativamente maior se as ferramentas combinadas forem capaz de atuar em conjunto, compartilhando informações e resultados de análise e desencadeando procedimentos automáticos de resposta de uma forma coordenada.

A falta de padronização e a limitada interoperabilidade entre os sistemas de detecção representam, nesse caso, um sério obstáculo. Uma iniciativa importante no sentido de superá-lo é o desenvolvimento do CIDF – Common Intrusion Detection Framework e da CISL - Common Intrusion Specification Language.

O CIDF estabelece padrões para a arquitetura dos IDS, definindo elementos, funções e estruturas genéricas. Conforme o CIDF, um sistema de detecção deve ser composto pelos seguintes elementos :

- **“ E ” boxes – representam os componentes capazes de observar eventos de interesse. Estão naturalmente associados aos agentes.**

- **“ A ” boxes – são os componentes encarregados de receber informações de eventos, conduzir análises e gerar relatórios. Correspondem às consoles de gerenciamento .**

- **“ D ” boxes – são componentes de banco de dados, responsáveis pelo armazenamento e análises avançadas das informações de eventos.**

- **“ R ” boxes – são componentes capazes de gerenciar o acionamento e a condução de procedimentos de resposta, em base automática ou manual.**

A CISL consiste numa linguagem empregada no intercâmbio de informações entre os componentes do IDS, dando suporte à comunicação e interpretação de informações brutas de evento, resultados de análise e instruções para procedimentos de resposta.

O desenvolvimento do CIDF e da CISL está a cargo do IDWG – Intrusion Detection Working Group, subordinado à IETF. Embora o padrão ainda não tenha sido homologado, espera-se que no futuro seu aperfeiçoamento e adoção pelo mercado leve ao desenvolvimento de uma geração de IDS interoperáveis e mais eficientes.

- **<u>Um novo modelo de Intrusion Detection System</u> :**

  A arquitetura de sistemas de detecção proposta no âmbito do CIDF aponta no sentido da modularidade como uma forma de organizar funções e permitir a interoperabilidade de sistemas distintos. A mesma característica é aproveitada como fundamento de uma idéia ainda mais inovadora : o desenvolvimento de sistemas de detecção constituídos por agentes autônomos.

  Os sistemas de detecção de agentes autônomos começam sua transição do campo experimental para o terreno ocupado por implementações funcionais. Essa evolução ganhou impulso recentemente pela liberação do AAFID – Autonomous Agents For Intrusion Detection - um sistema desenvolvido pelo COAST – Computer Operations Audit and Security Technology, da Purdue University.

O modelo proposto para os sistema de detecção de agentes autônomos baseia-se no emprego de um maior número de agentes, cada um deles funcionando com um alto nível de especialização e orientado a categorias específicas de eventos. Eles são capazes de estabelecer rotinas de cooperação com os outros agentes em ação no sistema, comunicando a estes a observação de eventos de interesse. Dependendo do grau de importância dos eventos observados, outros agentes autônomos com funções especializadas podem executar observações mais detalhadas ou tomar decisões que levem à notificação do evento. Esse método pretende otimizar o processo de análise, convergindo informações numa etapa anterior à notificação, o que contribui para um controle de erros mais eficaz.

É prevista também a mobilidade dos agentes, que podem ser redistribuídos automaticamente a partir de servidores de agentes, conforme as necessidades geradas pelos incidentes detectados. Técnicas avançadas como o emprego de algoritmos genéticos podem contribuir para a criação de agentes autônomos inteligentes, capazes de aprender com a observação de eventos ou a partir do resultado de análises e do acompanhamento de tendências.

Nesta linha de pensamento inserem-se os trabalhos conduzidos no Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo por Francisco Gomes Milagres e Mauro César Bernardes. Esses trabalhos avaliam possibilidades e especificam métodos para o uso de agentes autônomos em sistemas de detecção. O trabalho de Bernardes, particularmente, estabelece um modelo para comunicação e tomada de decisões entre agentes autônomos. Esse modelo é baseado em camadas conforme esquematizado na ilustração :

Por meio de mecanismos de troca de mensagens, os agentes que operam em uma camada podem informar agentes em uma camada superior, acionando funções específicas num grau crescente de atividade. Em certas circunstâncias, os agentes de uma camada superior podem acionar agentes da camada inferior, determinando a execução de tarefas de apoio ou complementares.

A proposta dos sistemas de detecção de agentes autônomos reduz ou elimina a necessidade de uma entidade central de gerenciamento, distribuindo suas responsabilidades entre os agentes especializados. Além do potencial aumento de eficiência que essa radical alteração pode trazer, outros benefícios são esperados. Eis alguns deles :

Flexibilidade – como o sistema abandona a estrutura integrada ou monolítica, adotada pela maioria dos IDS atuais, em favor de uma estrutura modular, fica facilitada a adequação às necessidades específicas de cada ambiente ou circunstância. A adição ou remoção de agentes ajusta as funções do sistema às exigências correntes.

Facilidade de manutenção – a divisão do sistema em módulos, relativamente simples se comparados aos sistemas integrados, reduz a complexidade das tarefas de manutenção a ele associadas. A configuração, substituição ou atualização de agentes tendem a ser mais fáceis e econômicas que as mesmas operações executadas sobre partes de um sistema integrado.

Escalabilidade - os sistemas de detecção de agentes autônomos oferecem maior escalabilidade que os sistemas integrados, porque a ativação de novos agentes e o estabelecimento de compatibilidade entre esses novos agentes e os agentes existentes é mais simples. Essa característica facilita o atendimento às necessidades de expansão e diversificação do sistema.

Tolerância a falhas – ao eliminar as entidades centrais de gerenciamento, os sistemas de detecção de agentes autônomos eliminam também um importante ponto único de falha, reduzindo a possibilidade de quebra da segurança como um todo a partir do comprometimento do módulo central do sistema.

4. <u>Protocolos seguros</u> :

Os protocolos seguros são componentes de software projetados para trabalhar em conjunto com os demais protocolos, provendo funções adicionais de segurança, como autenticação, verificação de integridade de dados e datagramas, manutenção da confidencialidade e outras.

- **<u>Aplicabilidade</u> :**

  O emprego de protocolos seguros depende fundamentalmente da camada em que esses protocolos irão atuar. Atuando na camada de Transporte ou de Aplicação, os protocolos seguros oferecem proteção aos dados, mas não são capazes de proteger as informações e parâmetros contidos nos cabeçalhos das camadas de Internet e Física. Isso implica numa limitação importante, uma vez que esses parâmetros afetam diretamente as comunicações em rede.

  O problema básico da aplicação de criptografia aos protocolos das camadas Física ou de Internet reside no fato de que isso obrigaria cada roteador ou dispositivo de conexão que opera nessas camadas a executar uma operação de decriptação, a fim de examinar o cabeçalho do datagrama ou frame, e em seguida uma operação de encriptação antes de encaminhá-lo. Além da sobrecarga extraordinária de processamento que essa prática viria impor aos dispositivos, possivelmente inviabilizando seu funcionamento, haveria ainda a vulnerabilidade representada pelo processamento em claro dos pacotes em cada dispositivo.

- **<u>Conceito de funcionamento</u> :**

  O funcionamento dos protocolos seguros envolve processos complexos, que podem variar caso a caso. Existem algumas técnicas e princípios, entretanto, que são

aplicadas em quase todas as implementações de protocolos seguros e podem esclarecer em termos genéricos o seu funcionamento. Faremos uma breve descrição de duas delas :

Encapsulamento – o encapsulamento consiste na reformatação de pacotes pela aplicação de um cabeçalho extra com novos parâmetros. O encapsulamento permite alterar a forma como os pacotes são processados nos dispositivos da rede sem que seu conteúdo original seja alterado. Ao alcançar seu destino, o cabeçalho do encapsulamento é removido e o pacote original é processado.

Criptografia – é um conjunto de métodos e processos que se destinam a alterar o formato de dados, tornando-os incompreensíveis. O processo de criptografia deve ser reversível, permitindo que os dados sejam reconstituídos ao seu formato original em condições controladas.

A criptografia emprega algoritmos e chaves. Os algoritmos são regras ou funções matemáticas para a alteração dos dados. As chaves são parâmetros ou variáveis utilizadas pelos algoritmos para criptografar os dados. Uma vez que os algoritmos dependem das chaves para criptografar, eles podem ser de conhecimento público, ficando o requisito de sigilo restrito às chaves.

Existem algoritmos padronizados de emprego muito difundido em técnicas criptográficas. São exemplos o RSA – Rivest – Shamir – Adleman , o DES – Data Encryption Standard e o Blowfish.

O número de bits que compõe a chave criptográfica impõe maior ou menor dificuldade à quebra e define o método de criptografia como forte ou fraco. Aceita-se que chaves com mais de 40 bits são necessárias para suportar criptografia forte.

O uso de chaves criptográficas associadas aos algoritmos permite que um destinatário escolhido, com conhecimento da chave usada para criptografar os dados, possa decriptografá-los, tendo acesso ao seu formato original. Esse processo

garante confidencialidade aos dados, evitando que destinatários não autorizados tenham acesso a eles.

Quando uma única chave é usada, diz-se que a criptografia é simétrica. Uma variação desse modelo é chamada de criptografia assimétrica. Os métodos de criptografia assimétrica empregam um par de chaves, uma delas pública e outra privada. As chaves pública e privada são geradas simultaneamente e guardam uma relação entre si. A chave pública é distribuída para os remetentes potenciais da mensagem. A chave privada permanece de posse exclusiva do destinatário. O remetente pode usar a chave pública de um destinatário escolhido para criptografar a mensagem. Essa mensagem só poderá ser decriptografada com a chave privada do destinatário. Isso garante que apenas esse destinatário terá acesso ao conteúdo da mensagem criptografada por esse método.

- **<u>Descrição de protocolos seguros</u> :**

Serão descritos aqui os seguintes protocolos :

a. **IPSec – IP Security**
b. **SSL – Secure Socket Layer**
c. **HTTP-S**
d. **SSH - Secure Shell**

a. **IPSec – IP Security :**

O IPSec é um multi-protocolo interoperável com o IP e outros protocolos. Ele foi projetado para fornecer integridade de dados, confidencialidade e autenticação sem afetar o funcionamento das redes e hosts que o utilizam O IPSec é um componente obrigatório do IPv6.

As técnicas de criptografia empregadas no IPSec não exigem algoritmos específicos. Existe um conjunto de algoritmos básicos recomendado para uso

em IPSec, mas cada desenvolvedor pode optar por utilizar em suas implementações o algoritmo que lhe parecer mais conveniente.

O IPSec é composto de três protocolos :

- **AH - Authentication Header**
- **ESP - Encapsulating Security Payload**
- **IKE – Internet Key Exchange**

O AH e o ESP são protocolos de autenticação bastante similares, apresentando entre si poucas diferenças, concentradas em variações de estrutura e no nível de criptografia utilizada por cada um deles.

É possível o emprego combinado do AH e do ESP em uma variedade de associações. Por esse meio podem ser implementadas configurações diversas, para suporte de comunicação segura host a host, estabelecimento de modalidades simples de VPN – Virtual Private Network e interação segura com firewalls.

Analisamos a seguir o funcionamento de cada um desses componentes do IPSec.

- **AH - Authentication Header :**

O AH provê integridade de dados e autenticação para datagramas IP. O AH deve ser aplicado a um datagrama antes de que esse datagrama seja fragmentado. Após aplicado, o datagrama IP pode ser fragmentado. O AH é descrito na RFC 1285.

O AH utiliza um cabeçalho que pode ser aplicado ao datagrama IP original de duas maneiras diferentes. Essas alternativas definem os dois modos de utilização do AH : o modo de transporte e o modo túnel.

No modo de transporte o cabeçalho IP do datagrama é removido, o cabeçalho AH é aplicado sobre a carga de dados ou payload do datagrama original, e um novo cabeçalho IP é composto. A estrutura do datagrama resultante é mostrada na figura abaixo :

CABEÇALHO IP	CABEÇALHO AH	DADOS ....

O modo de transporte protege os dados associados às camadas superiores, mas não todas as informações que fazem parte do cabeçalho IP. Os chamados campos mutáveis do cabeçalho IP não são autenticados. Ele exige menos processamento, mas apresenta um nível de segurança baixo, apresentando a desvantagem de não autenticar todos os campos do cabeçalho IP.

No modo túnel, o cabeçalho AH é aplicado sobre todo o datagrama, incluindo seu cabeçalho, conforme mostra a figura :

CABEÇALHO IP (NOVO)	CABEÇALHO AH	CABEÇALHO IP (ORIGINAL)	DADOS ...

O modo túnel exige maior processamento, mas fornece uma proteção mais efetiva, aplicada a todo o datagrama. Ele oferece adicionalmente a flexibilidade decorrente da possibilidade de serem empregados no cabeçalho novo endereços IP diferentes daqueles que constam do datagrama original. Isso permite o uso de sistemas intermediários denominados gateways seguros, para estabelecer comunicação segura entre si dispensando dessa tarefa os hosts de origem e destino.

- **ESP - Encapsulating Security Payload :**

O ESP fornece verificação de integridade, autenticação e criptografia dos datagramas IP. Como o AH, ele também fornece um serviço de proteção de

resposta a nível opcional. O ESP representa um método alternativo ao AH, com recursos mais amplos e poderosos.

A aplicação de um cabeçalho ESP a um datagrama envolve também a aplicação de um trailer e de dados de autenticação, posicionados após os dados. Conforme o modo de utilização, os dados podem incluir apenas a carga de dados do datagrama original ou o datagrama original em sua totalidade, inclusive com seu cabeçalho.

À semelhança do AH, o ESP pode também ser empregado em modo de transporte ou modo túnel, com as mesmas exigências e funcionalidades que esses modos oferecem ao AH.

No caso de emprego em modo de transporte, a estrutura do datagrama com o ESP aplicado seria a seguinte :

CABEÇALHO IP	CABEÇALHO ESP	DADOS ...	TRAILER ESP	AUTENT. ESP

Nesse modo o cabeçalho IP não é autenticado nem pode beneficiar-se das funções de encriptação. Contudo, o overhead gerado pelo seu emprego é menor do que em modo túnel.

No caso de emprego em modo túnel, o datagrama teria a estrutura ilustrada abaixo :

IP (NOVO)	ESP	IP (ORIGINAL)	DADOS ...	TRAILER ESP	AUTENT. ESP

O processamento extra exigido pelo emprego de ESP em modo túnel pode ser compensado, conforme as circunstâncias, pela proteção integral oferecida ao datagrama e pela possibilidade de utilização de endereços IP diferenciados.

- **IKE – Internet Key Exchange :**

O IKE deriva de dois outros protocolos, o ISAKMP – Internet Security Association and Key Management Protocol e o Oakley. Ele combina e aperfeiçoa as funcionalidades dos protocolos originais, dando suporte à geração automática de chaves criptográficas e à sua atualização. Por meio do IKE são trocadas as chaves criptográficas que serão usadas entre hosts envolvidos em uma transação segura. O estabelecimento de SA – Secure Associations - também pode ser feito por intermédio do IKE.

O IKE destaca-se principalmente pela possibilidade de executar as funções descritas com um alto nível de automatização, acelerando a execução de transações seguras.

As operações gerenciadas pelo IKE são organizadas em intercâmbios ou fases, de modo a otimizar o processo. Tarefas que requerem processamento intensivo são executadas num intercâmbio denominado Fase 1. Esse intercâmbio ocorre com uma frequência menor do que o segundo, denominado Fase 2, que engloba tarefas mais simples. Apesar da denominação sugerir uma sequência direta, a uma ocorrência da Fase 1 podem suceder-se diversas ocorrências da Fase 2.

b. SSL – Secure Socket Layer :

O Secure Socket Layer é um protocolo criado pela Netscape e de uso amplamente difundido em transações de e-commerce. Atualmente o SSL é um padrão aberto, incorporado a diversas implementações.

O SSL atua num nível superior da camada de Transporte, recebendo dados da camada de Aplicação e repassando-os ao protocolo de transporte, na maioria dos casos ao TCP.

O SSL utiliza um protocolo de handshake denominado SSL Handshake Protocol para estabelecer uma conexão segura baseada em estados de sessão. No estado inicial, denominado estado de sessão, o cliente e o servidor trocam uma sequência padronizada de mensagens com a finalidade de estabelecer parâmetros comuns que serão utilizados na criptografia dos dados. Nessa fase, informações como chaves públicas e algoritmos suportados são intercambiadas entre os hosts. É gerada também uma chave mestra de sessão, que será usada para criptografar os dados durante a fase de conexão.

Tendo sido concluídas com sucesso as tarefas referentes à fase de sessão, o SSL Handshake Protocol muda seu estado para conexão. Com a conexão estabelecida, a transmissão e recepção dos dados será gerenciada pelo SSL Record Protocol.

Na fase de conexão, o SSL Record Protocol executará a reformatação dos dados, de modo a facilitar as operações subsequentes. Ele poderá executar também uma compactação opcional. Os dados serão criptografados ou decriptografados com o uso da chave mestra e dos parâmetros definidos na fase de sessão. Algumas funções adicionais de autenticação também podem ser executadas.

c. HTTP – S :

O HTTP-S ou Secure HTTP é uma extensão do protocolo HTTP que oferece recursos como autenticação entre cliente e servidor, criptografia e controle de sessão. Ele consiste na verdade em um conjunto de cabeçalhos normalmente aplicáveis às mensagens HTTP e que podem ser usados também no controle dessas funções. Os cabeçalhos com conteúdos adaptados às funções de segurança encapsulam mensagens HTTP comuns, outras mensagens HTTP-S ou dados em estado bruto.

Os recursos de criptografia suportados pelo HTTP-S não são vinculados a algoritmos específicos. As implementações do protocolo podem utilizar-se de algoritmos não padronizados, o que lhes confere grande flexibilidade.

Enquanto o SSL foi projetado para estabelecer conexões seguras, o HTTP-S é usado para prover segurança a nível de mensagem. Os dois protocolos podem, inclusive, ser usados em conjunto.

d. SSH – Secure Shell

O SSH é constituído de um protocolo em torno do qual são agrupados módulos que desempenham funções de cliente e servidor. Cria-se assim um conjunto centrado no protocolo e que pode ser utilizado com finalidades diversas.

Os conjuntos SSH permitem conexões seguras a partir de um mecanismo de autenticação baseado em algoritmos eficientes como o RSA. O SSH criptografa toda a sessão, inclusive as senhas enviadas. Dessa forma, ele substitui com vantagem o utilitário rlogin, originalmente empregado em ambientes UNIX para estabelecer conexões com hosts remotos. Como o rlogin transmite as senhas em claro, ele oferece uma vulnerabilidade que pode ser eliminada pelo método implementado com o SSH.

Outra funcionalidade importante do SSH é a de permitir a execução de comandos remotamente. Essa funcionalidade justifica seu emprego como substituto de utilitários originais do UNIX, como o rsh e rexec, que carecem de mecanismos de autenticação confiáveis e são comprovadamente inseguros. A cópia e transferência de arquivos a partir de hosts remotos é outra tarefa suportada pelo SSH.

O SSH é um padrão comercial desenvolvido pela SSH Communications Security Ltd. Está disponível para ambientes UNIX e Windows NT, além de outros sistemas operacionais.

<u>Avaliação crítica</u> :

A aplicação de protocolos seguros constitui-se numa contramedida de eficiência comprovada em muitos casos. Como vantagens significativas do emprego desse método surgem a possibilidade de utilizá-lo com sucesso em um ambiente diferenciado e multi-plataforma e a concentração dos custos de sua utilização nas fases de desenvolvimento e implementação, sendo dessa forma uma contramedida pronta para o uso, que requer pouca ou nenhuma participação dos utilizadores.

Como principal obstáculo ao emprego dessa técnica estão as limitações associadas aos processos criptográficos, desde os requisitos de processamento e custos em termos de tempo que são por eles impostos até as questões relacionadas à produção e utilização de algoritmos.

Capítulo 5

Política de segurança

Política de segurança

A complexidade crescente dos ambientes informatizados e das facilidades oferecidas pela tecnologia exige que as medidas destinadas a apoiar a segurança dos dados e das comunicações sejam planejadas criteriosamente e desenvolvidas de forma coordenada.

A política de segurança é um conjunto de diretrizes, normas, regulamentos e instruções que formalizam métodos e condutas, definem procedimentos e estabelecem responsabilidades, contribuindo para a criação e manutenção de um ambiente organizado e seguro, onde o emprego de contramedidas de escopo tecnológico pode ser facilitado para alcançar resultados mais amplos e consistentes.

Assim composta, a política de segurança será usada como referência básica para :

- **Definir controles de utilização para serviços e aplicações.**
- **Estabelecer direitos e restrições de acesso a recursos.**
- **Definir medidas preventivas e corretivas das violações da segurança.**
- **Orientar análises de riscos.**
- **Conduzir processos de auditoria.**
- **Definir medidas preventivas e corretivas das violações da segurança.**

A política de segurança deve refletir as características da instituição ou empresa em benefício da qual foi criada, considerando, além das metas e objetivos, o ambiente institucional e a cultura da organização.

A elaboração de uma política de segurança deve ser orientada em níveis estratégico, tático e operacional. A cada um desses níveis corresponderá uma classe de disposições, com diferentes graus de exigência e detalhamento.

Assim, ao nível estratégico corresponderão diretrizes e princípios de aplicabilidade genérica, que contemplarão aspectos permanentes e comuns a todos os setores e atividades. As diretrizes de segurança para toda uma organização devem ser delineadas

nos mais altos escalões administrativos, a fim de garantir sua vinculação às necessidades mais gerais e aos objetivos permanentes da empresa.

Num nível tático, serão definidas normas de emprego mais específico, dependentes da estrutura organizacional e orientadas à aplicação dos princípios estabelecidos nas diretrizes. As normas devem ser relacionadas a ambientes e grupos específicos e podem ser direcionadas em separado ao pessoal técnico e aos usuários de informática em geral. No primeiro caso, elas tratarão de aspectos como política de senhas e métodos de backup. No segundo caso, estarão relacionadas às atividades dos usuários, abordando tópicos como utilização das senhas e organização de dados a cargo de cada setor ou funcionário.

Já a nível operacional, a política de segurança deve definir métodos, condutas e rotinas associadas a cada atividade ou situação específica. Neste caso, o nível de detalhamento é ainda mais alto, devendo atender às necessidades de informação imediatas que decorrem ou podem decorrer da execução de cada tarefa.

Na elaboração da política de segurança, as vulnerabilidades dos sistemas devem ser consideradas em conjunto, associando-se aquelas fundamentadas em características tecnológicas com as de natureza física ou as que são condicionadas por fatores humanos. A composição de um painel detalhado das vulnerabilidades existentes irá subsidiar a definição de procedimentos de segurança contextualmente eficazes, em condições de compensar o conjunto de vulnerabilidades e não apenas atender a necessidades isoladas.

Após sua elaboração, é necessário que a política de segurança da empresa seja apoiada por um processo de documentação e divulgação eficiente e torne-se objeto de um esforço de acompanhamento capaz de garantir as adaptações e atualizações necessárias a mantê-la sempre em condições de aplicabilidade plena.
Finalmente, a política de segurança deve contribuir para formar uma cultura em que a segurança do ambiente seja uma preocupação constante e contínua, elevando o nível de participação e conscientização de técnicos e usuários. Essa contribuição deve ser considerada desde a fase de elaboração da política até sua aplicação prática e cotidiana.

Capítulo 6

Conclusões

Conclusões

As análise e observações conduzidas no corpo deste trabalho permitem elaborar certas conclusões genéricas sobre as vulnerabilidades próprias do conjunto de protocolos TCP/IP.

O conjunto de protocolos TCP/IP tem seu emprego continuamente ampliado porque representa um padrão consensual e porque apresenta níveis de eficiência satisfatórios. Provendo interoperabilidade entre plataformas, poderosos recursos de conectividade e suporte a aplicações que fazem uso intensivo da comunicação em rede, como os sistemas voltados para e-commerce, o TCP/IP é capaz de atender às necessidades correntes do atual estado de desenvolvimento tecnológico e às exigências impostas pelas novas formas de comunicação e trabalho.

O protocolo IP, que provê as funcionalidades críticas no conjunto, é essencialmente inseguro. Sua forma de operação e características dão margem a explorações intencionais e formas variadas de uso indevido que podem afetar o funcionamento de serviços de rede ou comprometer requisitos básicos da segurança da informação, como sua integridade, confidencialidade e autenticidade.

As limitações de segurança que o protocolo IP revela decorrem basicamente das necessidades que definiram o cenário em que seu desenvolvimento teve início. Nesse cenário, foi previsto seu emprego num ambiente controlado e concedida prioridade às capacidades de comunicação em detrimento de processos abrangentes de verificação e segurança. Esses fatores produziram um protocolo flexível, mas sujeito a um grande número de falhas de segurança.

As vulnerabilidades do protocolo IP e dos demais protocolos que a ele são associados podem ser compensadas por aperfeiçoamentos no sentido de prover uma comunicação confiável, protegendo não só os dados transportados, mas também os parâmetros e informações de controle que orientam a comunicação. Técnicas como o encapsulamento de pacotes e a criptografia de dados e de parâmetros são exemplos de aperfeiçoamentos

que já foram incorporados com sucesso em implementações aperfeiçoadas do conjunto ou como módulos opcionais, associados a aplicações específicas. Um número significativo de alternativas puderam ser testadas, inclusive em campo, e muitas delas revelaram-se funcionais e seguras. O IPSec e o SSL exemplificam esse êxito e sinalizam um rumo para o desenvolvimento futuro.

O melhor caminho para tornar o IP um protocolo seguro, capaz de desempenhar suas funções de forma confiável, está na evolução de seu conceito e de suas especificações. A essa evolução devem seguir-se o desenvolvimento de implementações específicas e o surgimento de aplicações seguras, adequadas a novos formatos de emprego que já estão sendo delineados, como os Application Service Providers e as novas modalidades de ferramentas de groupware baseadas em redes peer-to-peer.

O IPv6 representa um resultado consistente do esforço empreendido para acrescentar funcionalidades mais amplas e níveis de segurança mais altos ao protocolo. A incorporação de técnicas de criptografia a essa nova versão deve contribuir para eliminar ou controlar algumas das vulnerabilidades mais preocupantes da versão atual do protocolo.

É esperado que a substituição do IPv4 pelo IPv6 seja feita ao longo de um processo demorado. A especificação cuidadosa de requisitos de interoperabilidade entre as versões do protocolo reflete essa perspectiva. Mesmo antes que os benefícios da utilização do IPv6 possam afetar a questão da segurança de forma ponderável, o emprego de contramedidas apropriadas representa uma alternativa de curto e médio prazos para amenizar as deficiências que decorrem do uso do TCP/IP.

Hoje temos contramedidas sofisticadas, que apoiam-se em técnicas e em implementações específicas para operar funções avançadas de gerenciamento e controle, sendo capazes de fornecer às comunicações um nível de segurança superior. Contudo, a simples utilização de uma dessas contramedidas não pode garantir segurança para o ambiente da rede IP. É recomendável coordenar a ação de várias delas num esforço integrado, como forma de complementar capacidades e compensar limitações.

Esse esforço integrado deve ser delineado a partir do estabelecimento de políticas de segurança adequadas às necessidades e características de cada empresa ou instituição. Riscos e vulnerabilidades devem ser avaliados e medidas preventivas ou corretivas devem ser selecionadas e implementadas não sob uma perspectiva exclusivamente técnica, mas em conformidade com os objetivos e necessidades previstos na política de segurança.

Bibliografia

Bibliografia

1. BELLOVIN, S.M.. **Security Problems in the TCP/IP Protocol Suite.** Disponível por www em http:// www.modulo.com.br.

2. FRANCISCO GOMES MILAGRES. A aliança dos agentes móveis e tecnologias contra os hackers. **Developers' Magazine**, Rio de Janeiro, RJ, ano 5, n.54, p.26-27, fev. 2001

3. **Frequently asked questions on PPTP.** Disponível por www em http://www.counterpane.com/pptp-faq.html

4. GONÇALVES, Marcus. **Firewalls :** Guia completo. Rio de Janeiro, RJ : Editora Ciência Moderna Ltda., 2000. ISBN: 85-7393-102-7.

5. HEYWOOD, Drew; SCRIMGER, Rob. **Networking with Microsoft TCP/IP Certified Administrator's Resource Edition**. Indianapolis, Indiana : New Riders Publishing, 1997. ISBN: 1-56205-791-X.

6. **Hypertext Transfer Protocol -- HTTP/1.0.** Disponível por www em http://www.ics.uci.edu/pub/ietf/http/rfc1945.html

7. **INTERNETWORKING with Microsoft TCP/IP on Windows NT 4.0** : Course Workbook : Microsoft Corporation,

8. LUIZ PAULO MAIA. Analisando ataques do tipo Distributed Deny of Service – DDoS. **Developers' Magazine**, Rio de Janeiro, RJ, ano 5, n.54, p.26-27, fev. 2001

9. MURHAMMER, Martin et al. **TCP/IP Tutorial e Técnico**. São Paulo, SP : Makron Books, 2000. ISBN: 85-346-1188-2.

10. NORTHCUTT, Stephen. **Como detectar invasão em rede :** Um guia para analistas. Rio de Janeiro, RJ : Editora Ciência Moderna Ltda., 2000. ISBN: 85-7393-070-5.

11. NOVELL intraNetWare Administration. Orem, Utah : Novell Inc., 1999.

12. NOVELL Networking Essentials Student's Handbook. Orem, Utah : Novell Inc., 1999.

13. RANGEL, Ricardo Pedreira. **Passado e Futuro da Era da Informação**. São Paulo : Nova Fronteira, 1999. ISBN85-209-0980-9

14. RUSSEL, Charlie; CRAWFORD, Sharon. **Microsoft Windows NT Server 4.0 :** Guia Autorizado Microsoft. São Paulo, SP : Makron Books, 2000. ISBN: 85-346-0811-3.

15. SECURE Web Acces with MS Proxy Server 2.0 : Course Workbook : Microsoft Corporation,

16. STRAUCH, Suzana Beatriz de Miranda. **Aspectos de Segurança no Protocolo IP**. Porto Alegre : PPGC da UFRGS, 1999. 100f.:il

17. TOWNSLEY, W. et al. **Layer Two Tunneling Protocol "L2TP"**. Disponível por FTP anônimo em isi.edu/in-notes/rfc2661.txt

18. WINDOWS NT 4.0 Server Resource Guide. São Paulo, SP : Makron Books, 1997. ISBN: 85-346-827-X.

Sites consultados :

Os seguintes sites forneceram dados e informações utilizados nesse trabalho :

- **http://packetstorm.securify.com**
- **http://rootshell.com**
- **http://support.microsoft.com**
- **http://www.atstake.com**
- **http://www.counterpane.com**
- **http://www.guninski.com**
- **http://www.ietf.org**
- **http://www.infosecuritymag.com**
- **http://www.linux.org**
- **http://www.mindsec.com**
- **http://www.nmrc.org**
- **http://www.novell.com**
- **http://www.ntbugtraq.com**
- **http://www.sans.org**
- **http://www.ssh.com**
- **http://www.terisa.com**