**Uma estrutura de riscos avançadas**

Em 2018, o NIST lançou a segunda revisão de sua publicação especial sobre riscos, SP 800-37, Risk Management Framework. A estrutura de gerenciamento de riscos apresenta um processo disciplinado, estruturado e flexível para gestão dos riscos à segurança e privacidade. Ela abrange a classificação de segurança da informação, seleção, implementação e avaliação de controles, autorizações para sistemas e controles comuns, e monitoramento contínuo de riscos. A estrutura de gerenciamento de riscos considera os riscos em três níveis: riscos dos sistemas de informação, risco à missão ou processo de negócios e riscos para a empresa como um todo. O processo de gerenciamento de riscos envolve a preparação do material sobre os riscos necessários para fazer um gerenciamento de riscos. Em seguida, um processo de seis fases: classificação, seleção, implementação, avaliação, autorização e monitoramento. A fase de preparação envolve sete ações, sendo três relacionadas a riscos e quatro associadas a controles. São elas: indicar pessoas para funções de gerenciamento de riscos, preparar o contexto do gerenciamento de riscos, também chamado de estratégia de riscos, concluir uma avaliação de riscos em toda a organização, definir linhas de base para controles conforme normas pertinentes para a organização, identificar controles comuns, priorizá-los com base no possível impacto de um ataque e desenvolver um plano para monitorar a eficácia dos controles. A fase de classificação se sobrepõe em parte à fase de preparação, pois exige uma análise completa dos sistemas de TI em uso, principalmente a identificação das características dos sistemas e das informações que processam e armazenam. A próxima etapa é determinar os níveis de impacto na confidencialidade, disponibilidade e integridade. O último passo é conseguir o endosso ou autorização da empresa para os três níveis de classificação de impacto. A fase de seleção exige que os controles sejam escolhidos e adaptados ao ambiente do sistema específico para mitigar todos os riscos ao sistema que extrapolem o apetite por riscos. Isso é avaliado determinando o nível de risco e identificando, com base no contexto de riscos, onde são necessários controles. Estas são as etapas da fase de seleção: escolha dos controles, seja com a adoção de um conjunto de controles de linha de base, um conjunto personalizado de controles orientados por avaliações de riscos ou uma combinação de ambos. Adequação dos controles ao ambiente operacional. Alocação dos controles aos sistemas, garantindo que os requisitos de negócios específicos para segurança naquele sistema sejam atendidos pelas pessoas, processos e tecnologia. Documentar os controles para cada sistema em um plano de segurança do sistema. Desenvolver e implementar a abordagem de monitoramento contínuo da eficácia dos controles. E conseguir a aprovação da empresa para os planos de segurança dos sistemas e o processo de monitoramento contínuo. A próxima fase é implementar os controles que foram identificados para o sistema e manter os planos de segurança do sistema em conformidade. A fase de avaliação envolve analisar o sistema ao longo de sua vida útil para assegurar que os controles sejam eficazes e não haja provas de violação. Essa fase do ciclo de vida do gerenciamento de riscos tem sete etapas. Seleção do avaliador, com base nas qualificações dos candidatos e no conhecimento do alvo, desenvolver o plano de avaliação, executar o plano de avaliação dos controles, relatar a eficácia dos controles apresentando conclusões e recomendações, corrigir as descobertas que possam ser resolvidas imediatamente, e desenvolver um plano de ação geral para as conclusões que não possam ser resolvidas de imediato. O objetivo da fase de autorização é a responsabilidade organizacional, exigindo que um gerente sênior determine se os riscos à segurança e privacidade representados pelo conjunto geral de atividades e planos de gerenciamento de riscos são aceitáveis. Essa fase tem cinco etapas: desenvolvimento da proposta, análise gerencial, resposta adicional para gerenciamento de riscos a quaisquer problemas levantados, decisões de aprovação para cada sistema e um relatório de autorização. A última fase é o monitoramento. É uma etapa essencial da estrutura que proporciona a conscientização situacional contínua sobre a postura de segurança e privacidade do sistema de informação e da organização em apoio às decisões de gerenciamento de riscos e inclui algumas das fases acima. Abrange sete etapas. O sistema e o ambiente são monitorados para observar possíveis alterações. São realizadas avaliações ao vivo, conforme necessário. Para qualquer problema identificado, é dada uma resposta. São mantidos documentos de gerenciamento de riscos. Os riscos à segurança e privacidade são relatados periodicamente. São concedidas autorizações aos sistemas, conforme necessário. E os sistemas são descartados com segurança quando deixam de ser necessários.

**Gerenciamento da segurança com o COBIT**

Uma das estruturas de TI mais importantes para a empresa é o COBIT, objetivos de controle para TI. O COBIT é publicado pela ISACA, Information Systems Audit and Control Association, e sua finalidade é garantir que as empresas tenham em vigor um conjunto eficaz e auditável de processos de governança e gerenciamento de TI que agregue valor para os interessados. O COBIT foi desenvolvido levando em conta um conjunto de processos agrupados em quatro áreas: planejamento, construção, entrega e monitoramento. No canto superior esquerdo, vemos o grupo de planejamento, conhecido na íntegra como alinhamento, planejamento e organização, com seus 13 processos de APO. Abaixo vem o grupo de construção, aquisição e implementação, com 10 processos. Na parte inferior do diagrama estão os seis processos do grupo de entrega, serviço e suporte e, à direita, o grupo de monitoramento, avaliação e análise, com três processos. A estrutura do COBIT é utilizada pelo setor financeiro para a realização de auditorias externas de controles gerais de TI. Consequentemente, ter uma estrutura de segurança alinhada ao COBIT é o primeiro passo para implementar um ambiente de TI que atenda às obrigações regulatórias. De uma perspectiva de cibersegurança, existem dois processos fundamentais: APO13, gerenciar a segurança, no grupo de planejamento, e DSS05, gerenciar serviços de segurança, no grupo de entrega. É claro que a segurança faz parte de muitos outros processos. Por exemplo: a gestão de incidentes de segurança é importante, mas se enquadra no processo geral DSS02, gerenciar solicitações de serviço e incidentes. Vamos ver o APO13, gerenciar a segurança, que define os requisitos para gestão da segurança. A descrição do processo é definir, operar e monitorar um sistema de gestão da segurança da informação. Ele tem cinco metas: apoiar a conformidade empresarial e de TI, apoiar o gerenciamento dos riscos empresariais e de TI, contribuir para a transparência dos custos e benefícios de TI, garantir a segurança das informações, infraestrutura e aplicativos, e disponibilizar informações confiáveis para a tomada de decisões. O APO13 consiste em três objetivos de controle. APO13.01: estabelecer e manter um sistema de gestão de segurança da informação empresarial. APO13.02: definir e gerenciar um plano de segurança que estabeleça um conjunto de objetivos para avançar rumo à postura de segurança desejada. E APO13.03: monitorar e revisar o SGSI. O sistema de gestão de informações corporativas, ou SGSI, define a abordagem adotada para garantir que a segurança da informação seja eficaz, e isso geralmente está alinhado ao conjunto de requisitos da norma internacional ISO 27001 para sistemas de gestão de segurança da informação. Embora o APO13.01 seja um objetivo de controle único, para atendê-lo, é preciso implementar vários controles de nível inferior da série de normas ISO 27000.

**O COBIT para segurança operacional**

Vamos analisar o segundo processo focado em segurança, DSS05, Gerenciar serviços de segurança. A descrição desse processo é proteger as informações da empresa para manter o nível de riscos à segurança da informação aceitável para a empresa conforme a política de segurança, estabelecer e manter funções de segurança da informação e privilégios de acesso e realizar monitoramento de segurança. Basicamente, o DSS05 define os requisitos de segurança operacional. O DSS05 apresenta processos operacionais para atender a três das cinco metas do APO13: apoiar a conformidade empresarial e de TI, apoiar o gerenciamento dos riscos empresariais e de TI, e garantir a segurança das informações, infraestrutura e aplicativos. Existem sete objetivos de controle de segurança operacional no DSS05 que servem de base para um programa defensivo de cibersegurança. Vamos analisá-los. O primeiro objetivo de controle é proteger contra malwares, um dos principais desafios atuais da cibersegurança. Para se proteger deles, há uma série de atividades e controles. Isso inclui antivírus, patches de segurança, conscientização sobre informações de segurança – em termos atuais, inteligência contra ameaças cibernéticas –, gestão de mudanças, filtragem de segurança de tráfego da web e e-mails, e treinamento em segurança. O segundo objetivo de controle é gerenciar a segurança de redes e conectividade. Isso inclui estabelecer e fazer cumprir a política de conexões de rede, imposição de entrada de senha, configuração e uso de firewalls e sistemas de detecção de intrusão, protocolos de segurança de rede e criptografia de comunicações, configuração de rede, mecanismos de segurança para garantir a transmissão e recebimento confiáveis, e testes de controles de segurança de rede e testes de penetração. Um controle crítico para esse objetivo é a segregação da rede. Gerenciar a segurança de pontos de extremidade abrange a segurança de notebooks, desktops, servidores, dispositivos móveis e equipamentos de rede. Isso exige a implementação de controles para garantir que os pontos de extremidade sejam configurados com segurança, protegidos para remover portas e protocolos desnecessários e que o acesso remoto seja gerenciado. O próximo objetivo de controle é gerenciar a identidade dos usuários e o acesso lógico. O gerenciamento de identidades e acesso é um objetivo de controle muito complexo e pode facilmente consumir metade do esforço em um programa de cibersegurança maduro. Também é a área que gera uma boa parte das conclusões de auditoria, por isso vale a pena se concentrar nela. Esse processo exige que as identidades sejam gerenciadas desde a criação até a remoção, os direitos de acesso sejam estabelecidos e mantidos conforme as funções e responsabilidades da organização, o acesso a sistemas e informações seja autorizado e autenticado, o acesso privilegiado seja controlado com rigor, os direitos de acesso sejam revisados com regularidade e trilhas de auditoria de acesso adequadas sejam mantidas. O quinto objetivo de controle aborda o gerenciamento do acesso físico às TIs. Isso inclui proteção do perímetro, como cercas, portas e fechaduras, sistemas de detecção de intrusos, controles de acesso para datacenters e escritórios, cartões de identificação e procedimentos de controle de visitantes. Cada vez mais, o uso de infraestruturas baseadas em nuvem está diminuindo o esforço necessário para gerenciar essa área, mas aumentando a dependência e a supervisão da segurança de terceiros. O gerenciamento de documentos confidenciais é um aspecto cada vez mais importante da segurança, pois o foco das medidas de proteção vai se estreitando desde o perímetro até as próprias informações. Com os funcionários levando notebooks da empresa para casa e enviando dados para dispositivos móveis, os dispositivos de segurança do perímetro, como firewalls corporativos, deixam de proteger as informações corporativas. É preciso empregar novas técnicas, como o gerenciamento de direitos digitais e a criptografia de e-mail para dispositivos móveis. Esse processo inclui ainda controles de pontos de extremidade, como senhas de computadores e notebooks, impressão segura e códigos PIN em dispositivos móveis. Por fim, monitorar a infraestrutura para eventos relacionados à segurança proporciona os controles de detecção necessários para identificar violações de segurança no caso de falha dos controles de prevenção da empresa. Esse objetivo de controle inclui a operação de sistemas de detecção e prevenção de intrusão, logs e alertas de eventos relacionados à segurança, operações de gerenciamento de logs, sistemas de monitoramento de informações e eventos de segurança, encaminhamento de incidentes de segurança ao processo de gerenciamento de incidentes, análise forense e gerenciamento de provas. Todas essas atividades são essenciais em um centro de operações de cibersegurança.

**Introdução aos controles de cibersegurança**

Cibersegurança significa proteger o que está no ciberespaço contra ataques. Fazemos isso usando controles de segurança adequados com eficácia comprovada. Além disso, os controles não podem inibir a capacidade da organização de fazer negócios e precisam ter boa relação custo-benefício. Como resultado, eles não costumam proporcionar uma mitigação total. Geralmente restará um nível baixo de riscos aceitáveis, conhecidos como riscos residuais. Podemos mostrar esse conceito da seguinte maneira: as ameaças, as vulnerabilidades e os impactos juntos formam os chamados riscos inerentes. Quando os controles são implementados, o que resta são os riscos residuais. Os controles podem mitigar os riscos de diferentes maneiras: dissuadindo, prevenindo ou detectando ataques e corrigindo os danos causados por um ataque. Ao projetar uma estratégia de controle, é importante combinar controles para proporcionar defesa em profundidade. Isso é conhecido como estratégia de controle multinível. Vamos ver um exemplo. Para proteger os dados contra modificações não autorizadas, primeiro aplicamos controles de acesso, para garantir que somente pessoas autorizadas acessem os dados. Caso isso falhe, monitoramos alterações nos dados. Se os dados foram alterados de forma mal-intencionada, poderemos recuperá-los com o backup.

**Estrutura de controle de cibersegurança**

Embora uma empresa possa aplicar controles como resposta personalizada aos riscos corporativos, em muitos casos uma autoridade externa orienta que um conjunto predefinido de controles seja adotado como linha de base para a segurança. Um exemplo de política governamental é a SP 800-53, Controles de segurança e privacidade para organizações e sistemas de informação federais. A autoridade pode ser um órgão do setor, como o PCI Security Standards Council, que exige que os comerciantes sigam a norma de segurança de dados do PCI. Há duas estruturas de controle importantes na cibersegurança. Uma é a SP 800-53, do NIST, e a outra é a ISO 27002. Ambas são estruturadas como um conjunto de categorias de controle, no qual existe uma série de controles específicos. Embora as categorias e os controles sejam diferentes em cada norma, é possível associá-los. Essas duas estruturas de controle são muito mencionadas em outros sistemas de segurança, em particular, a estrutura de segurança cibernética do NIST. Os controles da ISO 27002 são descritos em uma hierarquia de três níveis: categoria de segurança, objetivo de controle e controle. Veja um exemplo. Aqui observamos que o controle de acesso é a categoria principal, o controle de acesso ao sistema operacional é o objetivo de controle, e a identificação e autenticação de usuários é o controle. Os controles da NIST SP 800-53 são descritos em uma hierarquia de dois níveis. Neste exemplo, identidade e autenticação são a família de controle e identidade e autenticação (usuários organizacionais) são o controle. A descrição é muito semelhante à do controle da ISO 11.5.2. Na implementação de uma estrutura de controle, a primeira etapa importante é criar a declaração de aplicabilidade. Ela é o principal vínculo entre a avaliação de riscos e a seleção de controles. Sua finalidade é apresentar provas de que os controles das ferramentas foram avaliados. Os controles que não forem aplicáveis não serão implementados, e a justificativa para omiti-los é registrada na declaração de aplicabilidade. A elaboração de uma declaração de aplicabilidade clara é uma boa forma de diminuir o esforço necessário para atender e manter uma postura de segurança conforme e eficaz. Há uma série de considerações específicas sobre os controles. Controles comuns podem ser herdados por um ou mais sistemas, reduzindo os esforços de implantação e operação contínuos e os custos. Quando forem necessários controles específicos, mas ainda não estiverem presentes ou não puderem ser implementados, serão necessários controles de compensação, como verificações de amostras de autorizações manuais na ausência de um processo de autorização eletrônica. Depois da implementação de um controle, ele precisa de testes regulares, e isso deve fazer parte da rotina de qualquer programa de conformidade. Os testes de controles envolvem duas fases: testar a eficácia do projeto e a eficácia operacional. A eficácia do projeto é examinada verificando se o controle implementado atende aos requisitos originais do projeto. Por exemplo: um teste de projeto do controle ISO 11.5.2, autenticação da identificação de usuários, deve envolver a verificação de arquivos de configuração para confirmar se o acesso ao sistema exige a entrada de um identificador de usuário e se é necessária uma senha ou outra forma de autenticação antes de permitir acesso ao sistema. A eficácia operacional envolve testar o sistema e garantir que o controle mantenha a eficácia contra ataques. Por exemplo: um testador de penetração pode tentar uma injeção de SQL no campo identificador de usuário em um formulário de logon para ver se consegue acesso sem digitar credenciais válidas.

**A norma de boas práticas do ISF**

O Information Security Forum é uma organização independente sem fins lucrativos que reúne e compartilha a experiência em segurança da informação de muitas corporações. Um de seus principais produtos é a Standard of Good Practice, atualizada anualmente. Trata-se de uma norma de boas práticas sobre riscos e controles para gerenciar a cibersegurança, na qual o gerenciamento de riscos subjacentes é a IRAM, metodologia de avaliação de riscos da ISF. A IRAM tem três fases: avaliação do impacto na empresa, avaliação de ameaças e vulnerabilidades e seleção de controles. A norma de boas práticas do ISF de 2018 estrutura seus controles em 17 categorias, 34 áreas e 131 temas. Ela é condizente com as principais normas de segurança da informação reconhecidas, como a ISO 27002, a estrutura de cibersegurança do NIST, o COBIT e o PCI DSS. Também se alinha com os controles necessários para atender ao Regulamento Geral sobre a Proteção de Dados da Europa. Vamos analisar um típico controle do ISF, monitoramento e melhoria da segurança. Ele tem duas áreas e oito temas. As duas áreas são auditoria de segurança, com cinco temas, e melhoria da segurança, com três temas. Se detalharmos o tema S12.1, monitoramento da segurança, ele tem um princípio, “a condição da segurança da informação da organização deve ser monitorada regularmente e relatada à diretoria executiva”, e um objetivo, “apresentar à diretoria executiva uma avaliação precisa, abrangente e coerente da condição da segurança da informação da organização”. Esse objetivo então se desdobra em vários subtemas, assim como os controles da ISO e do NIST. A Standard of Good Practice é uma abordagem setorial abrangente para a segurança, que disponibiliza um conjunto de controles para atender à maioria dos requisitos de conformidade.

**A ASD e seus principais controles**

As associações de normas da Austrália e da Nova Zelândia têm um forte histórico de contribuição para as normas internacionais. Normas como a ISO 31000 – Gestão de Riscos tiveram origem na AS/NZS 4360 Handbook of Risk Management. A Australian Signals Directorate, equivalente australiana da Agência Nacional de Segurança americana, publicou suas próprias normas de segurança. A principal referência é o Information Security Manual, que lista mais de 650 controles de segurança. Embora o ISM seja abrangente, é detalhado demais para ser uma referência de segurança útil fora das forças armadas. Para tentar torná-lo útil, em 2010 a ASD publicou seu Top 35, um conjunto gerenciável de controles de segurança menos abrangente que a ISO 27000, porém menos complicado que o ISM. Esses controles são classificados em cinco níveis de importância: essencial, excelente, muito bom, bom e limitado. Os controles são agrupados conforme a função: prevenir incidentes, limitar a magnitude de incidentes, detectar e responder a incidentes, recuperar-se de incidentes e evitar que pessoas mal-intencionadas tenham informações privilegiadas, ou seja, gestão de pessoal. Depois dessa publicação, a ASD promoveu vários subconjuntos desses controles. O primeiro foi os Top 4 da ASD, que, segundo a agência, são responsáveis por mais de 85% das violações. Nesse momento, houve um interesse internacional na abordagem da ASD. O Top 4 consiste na lista de permissões de aplicativos, aplicação de patches de aplicativos e sistemas operacionais e limitação de privilégios administrativos. A ASD também promove os Essential 8, oito controles essenciais que ampliam os Top 4 com configurações de macros, proteção de aplicativos, autenticação multifator e backups. Infelizmente, os controles da ASD têm dois pontos fracos importantes. Eles se preocupam em controlar o que usuários legítimos podem fazer, em vez de verificar como os invasores atuam, e a aplicação deles no ambiente empresarial nem sempre é fácil ou prática. A lista de permissões de aplicativos é um controle poderoso, principalmente em estações de trabalho, mas verificou-se que restringe as empresas e raramente é usada. Mesmo tarefas comuns, como participar de webconferências, podem se tornar difíceis quando se usa uma lista de permissões de aplicativos. Além disso, um ataque geralmente injeta um malware em um processo. Ele não precisa ser executado como aplicativo, então facilmente evita esse controle. Restringir privilégios administrativos é um bom controle, mas não resolve a questão da elevação de privilégios para técnicas de ataque generalizado.

**Como proteger os dados de cartões de pagamento**

Os cibercriminosos logo perceberam que os cartões de crédito eram um alvo atraente para ataques. O conselho do PCI, Payment Card Industry, também entendeu essa ameaça e publicou a Data Security Standard como estrutura prática para elaborar um regime de segurança robusto para os dados de titulares de cartões. Recentemente, diante de ataques patrocinados por Estados a dados pessoais e governamentais identificados nos vazamentos de Snowden, autoridades reguladoras governamentais promulgaram requisitos regulatórios para notificação de violações de dados, em particular, o Regulamento Geral sobre a Proteção de Dados da União Europeia. Isso aumentou a responsabilidade das empresas em caso de violação de dados, e agora é importantíssimo que qualquer empresa que receba pagamentos por cartão de crédito proteja suas informações e transações. Ao analisar a Data Security Standard, é útil entender a terminologia do PCI. Vamos ver os principais termos. Um comerciante é alguém que aceita cartão de crédito ou débito como forma de pagamento. Um prestador de serviços é alguém que fornece um serviço usado para armazenamento de informações de cartões de pagamento ou transações de um comerciante. Um avaliador de segurança qualificado é uma pessoa independente certificada para relatar a conformidade com o PCI. Um avaliador interno de segurança trabalha para o comerciante e é certificado para apresentar uma autoavaliação. Uma violação de dados é uma falha de segurança que resulta na perda de dados de titulares de cartões. Os dados do titular do cartão são o número da conta principal, o nome do titular, a validade e o código de serviço. Os dados de autenticação confidenciais são os dados codificados na tarja magnética ou chip, ou seja, o código de segurança do cartão e a senha. Em caso de violação de dados, a operadora do cartão iniciará uma investigação para determinar a causa. Se a empresa tiver certificação de conformidade com o PCI, contará com proteção jurídica. Caso contrário, poderá levar uma multa pesada ou perder o direito de aceitar cartões de crédito. Independentemente disso, enfrentará custos que podem incluir a substituição do cartão e talvez indenização do cliente. Nesse contexto, vamos analisar a Data Security Standard. A norma faz parte de um conjunto publicado pelo conselho do PCI e apresenta controles práticos e procedimentos de teste que definem com clareza o que precisa ser feito para alcançar a conformidade. A versão 3.2 da norma apresenta 12 áreas de requisitos técnicos e operacionais que abrangem quase 200 controles obrigatórios. O primeiro requisito é ter uma configuração de firewall eficaz. Isso significa definir normas para a configuração do firewall, testar todas as alterações do firewall e revisar as regras a cada seis meses. É necessário manter um diagrama de rede sempre que parte da rede armazene, transmita ou faça interface com dados de cartões de pagamento, e os fluxos de dados que atravessam a rede precisam ser definidos. O tráfego não relacionado a transações de titulares de cartões não pode ter acesso aos sistemas de titulares de cartões. Por isso, é normal ter uma zona de PCI segregada na rede interna, para que o tráfego possa ser gerenciado no próprio gateway do PCI. É necessária uma zona desmilitarizada para qualquer sistema com acesso direto à internet, e isso precisa ser protegido por firewall nos gateways internos e externos. Os firewalls não são apenas para empresas. Qualquer dispositivo móvel, inclusive de colaboradores, que tenha permissão para se conectar com a rede interna, precisa ter um software de firewall pessoal instalado e em funcionamento com uma configuração que não possa ser alterada pelo colaborador. Essa é uma consideração essencial ao pensar em ambientes BYOD, ou traga seu próprio dispositivo. O próximo requisito é que todas as senhas padrão e configurações inseguras sejam alteradas. Para permitir uma proteção eficaz, são necessárias configurações de segurança em cada dispositivo e componente do sistema, e todas as portas e serviços desnecessários devem ser removidos. Deve-se armazenar o mínimo possível e proteger os dados de titulares de cartões, e nenhum dado de autenticação confidencial pode ser armazenado, mesmo criptografado, após a conclusão da autenticação. O número da conta, quando exibido, deve ser mascarado, geralmente substituindo todos os dígitos por asteriscos, com exceção dos quatro últimos. Quando armazenado, o número da conta precisa ser protegido por criptografia forte ou hash unidirecional. O gerenciamento de chaves é uma parte fundamental de qualquer solução de criptografia e precisa implementado com eficácia. Os dados de titulares de cartões transmitidos em redes abertas, como a internet e redes sem fio desprotegidas, precisam ser criptografados, e nunca devem ser usados sistemas do usuário final, como e-mail e mensagens, para enviar números de contas desprotegidos. O esquema de criptografia precisa ser eficaz, e a eficácia pode mudar com o tempo. Por muitos anos, o protocolo SSL foi uma solução criptográfica comum para acesso via web. Em 2013, uma vulnerabilidade fundamental no esquema foi detectada e explorada no bug Heartbleed. Em seguida, o conselho de normas de segurança do PCI determinou que o protocolo SSL não era mais uma solução aceitável para a proteção dos dados de titulares de cartões. Os sistemas que processam dados de titulares de cartões precisam implementar um software antivírus para oferecer proteção contra malwares em servidores e dispositivos em pontos de extremidade. Como alguns malwares podem entrar em um sistema antes que sua assinatura seja incluída no banco de dados dos antivírus, também é preciso fazer varreduras periódicas. Ameaças e vulnerabilidades devem ser monitoradas por meio de alertas de fornecedores e feeds de inteligência contra ameaças, e patches críticos de segurança precisam ser instalados dentro de um mês após o lançamento. As contas de desenvolvimento e teste têm que ser removidas antes que os sistemas sejam colocados em produção, e o desenvolvimento personalizado precisa incluir a revisão do código-fonte antes da implementação, com atenção especial a vulnerabilidades comuns, como injeção de SQL e scripts entre sites. Os números de conta em produção não devem ser usados para testes. Os próximos seis requisitos da norma de segurança de dados do PCI tratam do nível de segurança exigido fora do ambiente PCI. Um requisito interessante é a restrição do acesso físico, que se estende a dispositivos para fins específicos usados para ler cartões. Caixas eletrônicos e, mais recentemente, leitores incorporados em equipamentos como bombas de combustível, são alvos constantes de criminosos que instalam skimmers, ou “chupa-cabras”, capazes de copiar dados de cartões de crédito. São esquemas grandes. Em 2015, uma varredura em 6.000 postos de combustível da Flórida encontrou 81 chupa-cabras em bombas de combustível. Estima-se que os criminosos faturem até US$ 3 bilhões por ano com esse golpe nos EUA.

**Problemas na nuvem**

A tecnologia de nuvem não é mais uma abordagem inovadora para a implantação de infraestrutura, mas sim uma tendência predominante. As soluções em nuvem podem ser implantadas pela equipe de TI da empresa ou por grupos comerciais. Existem três formas comuns de implantação de nuvem: infraestrutura como serviço, plataforma como serviço e software como serviço. Porém, existem muitas outras formas especializadas de serviços em nuvem. Em todos os casos, é necessário usar controles de segurança para proteger a solução em nuvem, assim como nas soluções locais. Contudo, existem diferenças nos controles ao usá-los em nuvem, e é preciso considerar alguns controles novos. A International Standards Organization elaborou a norma ISO 27017 para nuvem. Ela se baseia na ISO 27002 e inclui seis controles a mais. O NIST criou a SP 800-144 – Guidelines on Security and Privacy in Public Cloud Computing, que faz referência aos controles da SP 800-053. No entanto, a principal referência para segurança em nuvem é o Security Guidance for Critical Areas of Focus in Cloud Computing, da Cloud Security Alliance. Ela pode ser baixada aqui no site da CSA. A diretriz da Cloud Security Alliance é apresentada em 14 domínios, conforme mostrado. Esses domínios apresentam uma descrição completa da nuvem, suas necessidades de segurança e os controles necessários para proteger as implantações na nuvem. Essa diretriz é apoiada por uma Matriz de Controles em Nuvem, que também pode ser baixada. Ela apresenta os princípios fundamentais de segurança que devem ser adotados pelos fornecedores de nuvem e que podem ajudar os clientes a avaliar os riscos gerais de segurança de um provedor de nuvem. A matriz mostra com clareza as responsabilidades compartilhadas entre o provedor de serviços em nuvem e o cliente de acordo com a forma de nuvem. O mais importante é que ela faz uma referência cruzada dos controles com outras normas de segurança reconhecidas do setor. A versão 3.0.1 da CCM tem 133 controles em 16 domínios, que podem ser vistos aqui. Eles não se alinham com os domínios do Security Guidance for Critical Areas of Focus in Cloud Computing, mas fazem uma cobertura abrangente da segurança em toda a nuvem, começando pela segurança de aplicativos e interfaces e terminando com o gerenciamento de ameaças e vulnerabilidades. A Matriz de Controles em Nuvem é disponibilizada como uma planilha com mais de 60 colunas de informações relacionadas aos controles. O primeiro controle que vemos é o AIS-01, no domínio segurança de aplicativos e interfaces. Resumindo, esse controle exige que as interfaces de programação de aplicativos sejam baseadas em normas do setor. Veja que isso se aplica à computação em nuvem, armazenamento, aplicativos e dados. Também se aplica a software como serviço, plataforma como serviço e infraestrutura como serviço. Considerando que isso faz parte de um serviço implantado, é um controle que precisa ser implementado pelo provedor de serviços, e não pelo cliente. O AIS-01 tem grande relevância e está bem representado nas normas atuais. Se observarmos o IVS-10 – Segurança de máquinas virtuais – proteção de dados, vemos que isso é pertinente apenas para computação, mas é importante para os três serviços. Para quem leva a nuvem a sério, a diretriz e a matriz de controles da CSA são leituras essenciais.

**A proteção de dispositivos na internet**

Internet das coisas é um termo que significa tudo o que está conectado à internet que não seja um notebook, uma estação de trabalho ou um servidor comum. Uma definição de internet das coisas é: a interconexão pela internet de dispositivos de computação incorporados em objetos do dia a dia, permitindo que enviem e recebam dados. A Wikipédia dá mais detalhes, definindo a internet das coisas como um sistema de dispositivos de computação, máquinas mecânicas e digitais, objetos, animais ou pessoas inter-relacionados que recebem identificadores exclusivos e a capacidade de transferir dados por uma rede sem a necessidade de interação humano-humano ou humano-computador. A característica de uma coisa da internet é estar conectada com a internet. Ela pode apenas enviar dados, apenas receber ou enviar e receber. Uma classe importante são as coisas de baixo consumo de energia – aqueles objetos que têm bateria integrada e não têm fonte de alimentação externa. Muitas vezes, precisam ter vida útil de dez anos ou mais, portanto exigem operação com baixa potência. Uma das primeiras organizações a dar orientações sobre segurança para a internet das coisas foi a IoT Alliance Australia, integrante da Australian Communications Alliance. Sua diretriz inicial de segurança da internet das coisas foi publicada em 2017 e apresenta uma introdução à IoT e seus principais setores. Ela aborda questões legais, de privacidade, segurança, resiliência e sobrevivência, bem como considerações sobre o desenvolvimento de dispositivos IoT. Não há um conjunto definitivo de controles de segurança para a IoT, embora organizações como a OWASP e a GSMA tenham fornecido orientações. A IoT Security Foundation publicou o conjunto mais abrangente de controles em sua diretriz de segurança, propondo 142 controles agrupados em 13 áreas de conformidade, que podem ser vistas aqui. Um atributo adicional interessante que a IoT Security Foundation marcou para cada controle é sua classe de conformidade, distribuída em cinco valores relacionados aos dados gerados ou ao nível de controle proporcionado pelo dispositivo. O controle é pertinente para o dispositivo IoT quando sua classe de conformidade é igual ou superior à marca de controle. Classe 0 significa que um comprometimento provavelmente resultaria em pouco impacto discernível sobre um indivíduo ou organização. Classe 1 significa que um comprometimento provavelmente teria um impacto limitado sobre um indivíduo ou organização. Os dispositivos de classe 2 são projetados para resistir a ataques à disponibilidade que podem causar um impacto considerável sobre indivíduos ou organizações. Os dispositivos de classe 3, além disso, são projetados para proteger dados confidenciais. E os dispositivos de classe 4 podem afetar a infraestrutura crítica ou causar danos pessoais. Provavelmente os controles de segurança da IoT receberão muito mais atenção à medida que forem implantados em setores essenciais, como transportes e cidades inteligentes.

**Como garantir a eficácia da segurança**

Embora possamos projetar e implementar segurança em nossos sistemas, isso não garante que seja eficaz. Na verdade, a maioria das violações ocorre em soluções que teoricamente usam as boas práticas de segurança do setor. Os motivos das falhas de segurança variam. Algumas violações ocorrem devido ao projeto ou implementação insatisfatória. Algumas ocorrem devido à falta de higiene de segurança na empresa. E outras porque foram descobertas novas explorações que permitem que um invasor burle a segurança. A garantia é um processo que começa antes do lançamento de um sistema e continua ao longo da vida. Existem diversos métodos básicos de garantia que são usados pelas organizações para verificar seus sistemas. Os principais são: varreduras de vulnerabilidades, que identificam os serviços que podem ser acessados e verificam se todos os patches conhecidos foram aplicados; auditorias de controle de segurança, que predefinem um conjunto de condições com base no controle e preveem os resultados esperados; e testes de penetração, que usam explorações conhecidas e técnicas de exploração para tentar burlar o controle. Além dos próprios testes, as organizações podem contratar uma avaliação independente de seus sistemas. Isso costuma ser feito para certificar de alguma forma que os usuários podem confiar no sistema. Muitas organizações têm obtido a certificação ISO 27000 para toda a empresa, visando demonstrar que seus processos que abrangem o projeto, a implementação e a operação de sistemas são seguros. Muitas vezes, os fornecedores enviam seus produtos para avaliação de acordo com o chamado esquema de Critérios Comuns. Isso proporciona uma visão independente da garantia do produto, desde uma declaração básica de conformidade da documentação até o nível mais rigoroso de avaliação. O esquema de Critérios Comuns é reconhecido por muitos países e é o principal sistema usado para garantia de produtos. A garantia não é uma resposta perfeita. Ocorrem violações nas organizações certificadas, e os produtos avaliados também são explorados. Porém, ela aumenta consideravelmente o nível de dificuldade e evita que as organizações sejam vítimas de ataques simples.

**Série Rainbow**

* **Orange book – Computadores**
* **Red book – Redes**
* **Green book – senhas**

**Orange Book: conceitos iniciais sobre segurança de computadores**

A era da computação confiável começou na década de 1980 com a publicação de uma série de livros sobre os requisitos de segurança de TI para o Ministério da Defesa dos Estados Unidos. Esses livros ficaram conhecidos como série Rainbow. O livro mais conhecido é o Orange Book, que descreve o projeto de segurança de um computador capaz de lidar com informações confidenciais e não confidenciais, conhecido como computador confiável ou com segurança multiníveis. O Orange Book serviu de paradigma de segurança da informação na década seguinte. Até hoje a abordagem continua relevante e pode ser vista no raciocínio recente dos EUA sobre espaços confiáveis personalizados. O Orange Book estabelece os critérios técnicos necessários para o projeto de segurança e posterior avaliação de segurança do hardware, firmware e softwares do computador. Ele apresenta quatro conceitos essenciais na segurança da informação. Um monitor de referência que intermedia o acesso aos recursos do sistema, um modelo de segurança formal para leitura e gravação de informações, a ideia de uma base de computação confiável como subsistema contendo todo o código de segurança e os testes necessários para alcançar vários níveis de garantia. O conceito de monitor de referência é essencial em qualquer sistema que disponibilize controles e centros de computação com segurança multiníveis. O monitor de referência impõe controles de acesso entre sujeitos e objetos do sistema. O sujeito pode ser um usuário ou um módulo de programa, e o objeto pode ser um arquivo de dados ou uma função restrita do sistema. Vou usar o termo mais comum, “usuário”, em vez de sujeito. O monitor de referência tem três características essenciais. O mecanismo de validação de referência precisa ser inviolável, sempre deve ser invocado e tem que ser pequeno o bastante para ser analisado e testado por completo. O Orange Book introduz o modelo de Bell-LaPadula para gestão de fluxos de informações multiníveis. Usando esse sistema, o livro apresenta duas abordagens de segurança. O controle de acesso discricionário é usado para aplicar segurança na mesma classificação de informações visando restringir o acesso a informações com base em que realmente precisa acessá-las. Para isso, um administrador precisa manter uma lista de controle de acesso indicando a quais objetos cada usuário tem acesso. É o sistema normal de controle de pastas e arquivos que usamos hoje. O controle de acesso obrigatório é o modelo de Bell-LaPadula no qual cada usuário tem um certo nível de autorização ou direitos de acesso e um objeto é rotulado com um determinado nível de confidencialidade. Os rótulos de segurança que definem os níveis de confidencialidade no Orange Book são “restrito”, “confidencial”, “secreto” e “ultrassecreto”. O controle de acesso obrigatório tem duas regras. A primeira é a regra de segurança simples, que afirma que um usuário com um determinado nível de autorização não pode ler nada rotulado com um nível de confidencialidade maior, ao qual, por definição, não tem acesso. A segunda é a regra de segurança do asterisco, que afirma que um usuário com um determinado nível de autorização não pode gravar informações em um arquivo rotulado com um nível inferior, já que isso pode expor informações confidenciais a usuários não autorizados a acessá-las. A essência de um sistema de computador confiável é sua base de computação confiável, que contém todos os elementos do sistema responsáveis por dar suporte à política de segurança e ao isolamento de objetos nos quais a proteção se baseia. Os limites da TCB são chamados de perímetro de segurança. A TCB inclui hardware, firmware e software críticos para proteção. Ela precisa ser projetada e implementada de modo que nada fora da base de computação confiável seja confidencial ou pertinente para a gestão da segurança. Uma TCB deve ser o mais simples possível e condizente com as funções que deve executar para possibilitar testes adequados. A última parte, provavelmente a mais importante do Orange Book, é o sistema de classificação usado para avaliar a garantia dos sistemas. Resumindo, o sistema estabelece quatro níveis de garantia. Em cada nível, há uma ou mais camadas. Os níveis são D1 – proteção mínima, C1 a C2 – proteção discricionária, B1 a B3 – proteção obrigatória, e A1 – proteção verificada. Vamos analisar melhor os níveis de garantia mais adiante no curso.

**Cyber Kill Chain: a cadeia de destruição cibernética**

Na década de 1990, os ciberataques costumavam ser associados a brincadeiras de adolescentes entediados. Porém, a possibilidade de crimes pela internet ou de explorar a conectividade para levantamento de informações de inteligência não passou despercebida. Atualmente, os ciberataques vêm principalmente de criminosos organizados e agentes patrocinados pelo Estado, usando processos de negócios de ponta a ponta claramente definidos. Em 2009, a equipe de resposta a emergências cibernéticas da Lockheed Martin produziu um artigo inspirador sobre ataques cibernéticos chamado Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains, que pode ser baixado aqui no site da empresa. A pesquisa introduziu o conceito de Cyber Kill Chain, ou cadeia de destruição cibernética, que considera um ataque em sete etapas: reconhecimento, armamento, distribuição, exploração, instalação, comando e controle, e ação. Nem sempre um ataque avança passo a passo. Muitas vezes se sobrepõem, mas cada estágio representa um marco na execução do ataque. Reconhecimento significa encontrar um alvo e compreender suas características. As pessoas normalmente têm um endereço na internet que foi alocado pelo seu provedor de serviços de internet, ao passo que uma empresa pode ter vários endereços em seu domínio na internet. Um ciberataque contra um alvo corporativo começa com o endereço de um site conhecido e, em seguida, analisando o espaço da internet ao redor desse endereço, em busca de outros sistemas usados pelo alvo. A empresa vê isso como uma verificação de resposta em cada host de seu domínio. Isso é conhecido como varredura de endereços IP. Quando o invasor chega a uma lista de hosts ativos, ele analisa cada host para descobrir quais pontos de entrada estão expostos. Isso é conhecido como varredura de portas. Serve para identificar possíveis vetores de ataque e verificar as versões de software usadas nesses vetores. Hoje em dia, os ataques não são feitos manualmente. Um invasor costuma comprar tempo em uma rede de computadores comprometidos para executar varreduras automatizadas. Essas redes são chamadas de botnets e podem consistir em centenas de milhares, ou até milhões, de computadores comprometidos. Isso permite que o ciberataque seja executado em escala. Armamento significa pegar uma vulnerabilidade conhecida e personalizá-la a um alvo ou grupo de alvos específico e integrá-la para permitir a execução em uma plataforma automatizada de ciberataque. O malware armado pode ser desenvolvido para explorar uma vulnerabilidade de uma versão específica de um sistema operacional ou visar a um site específico de banco online. Na era do hacking como negócio, os cibercriminosos costumam adquirir malwares armados de desenvolvedores dedicados em vez de criar a própria versão. A forma mais comum de distribuir malwares é anexar um documento infectado – uma imagem em PDF ou outro item eletrônico – de modo que, quando o documento é aberto, o malware se instala automaticamente. Esse arquivo pode ser enviado para a vítima por e-mail, processo chamado de phishing. Outro caminho pode ser encontrar um site vulnerável infectado com o malware e enviar um convite por e-mail para que o alvo acesse o site. Quando a vítima visita o site, o malware é baixado, infectando sua estação de trabalho. Uma terceira maneira é usar senhas e IDs padrão integrados ao software no sistema-alvo ou usar uma senha e ID de usuário roubados para entrar no sistema-alvo e implantar o malware diretamente. Também é possível encontrar falhas em softwares expostos à internet e distribuir o malware manualmente. Na prática, um ataque normalmente exige que se conquiste uma “cabeça-de-praia” em um host exposto à internet e, em seguida, usá-la para penetrar mais no sistema e chegar ao verdadeiro alvo, que talvez não esteja conectado diretamente à internet. Por fim, é possível usar um pen drive infectado para distribuir o malware, o que pode ser muito eficaz se o sistema-alvo não estiver conectado à internet. Para isso, o usuário do sistema-alvo precisa ser convencido ou enganado a usar o pen drive. Nos ataques em anexo de e-mail e pen drive, o item infectado explora uma vulnerabilidade no software-alvo após a distribuição, quando o documento é aberto. Da mesma forma, um site comprometido pode baixar um código HTML que tira proveito de uma vulnerabilidade do navegador. No caso do acesso remoto, a fase de exploração pode usar a transmissão de um pacote para explorar uma vulnerabilidade no protocolo de um serviço exposto à internet, ou pode simplesmente usar credenciais craqueadas ou roubadas. Depois da etapa de exploração, o malware ou o invasor pode simplesmente agir, indo direto para a última fase da cadeia de destruição cibernética. No entanto, o mais comum é a instalação de um payload, ou carga de ataque, na memória ou disco rígido do sistema-alvo. Além disso, pode ser introduzido algum mecanismo para garantir que o payload seja reiniciado sempre que o sistema for reinicializado. Uma maneira de fazer isso no Windows é adicionar uma entrada de registro para executar automaticamente o payload quando o sistema for inicializado. Muitas vezes, o payload inclui um meio de manter o acesso contínuo em um comando show. Como o comprometimento do sistema costuma ser automatizado, assim que o payload é instalado, sua primeira ação é se conectar novamente ao servidor de comando e controle para se registrar como host comprometido. O invasor deve então enviar de volta um comando para que alguma ação seja tomada – listar os subdiretórios e arquivos, extrair arquivos com nomes específicos, modificar ou substituir softwares e assim por diante. Uma característica importante do payload é que ele pode determinar os endereços do servidor de comando e controle, que talvez mudem com o tempo. O mecanismo de ação exato do payload ao alcançar o alvo depende dos motivos do invasor. Um hacktivista talvez queira desfigurar um site. Um agente patrocinado pelo Estado pode querer roubar informações confidenciais. Já um cibercriminoso talvez queira acessar uma conta bancária para roubar dinheiro. No entanto, o que eles têm em comum é que, seja qual for a ação, é improvável que seja do interesse da vítima.

**Uma revisão das ameaças tradicionais**

Nos primórdios da computação, as ameaças de segurança enfrentadas pelas empresas refletiam as fraudes tradicionais antes da TI. Um exemplo de fraude tradicional são funcionários fantasmas na folha de pagamento, para retirar dinheiro e depositá-lo na conta de outra pessoa. Em 2012, uma mulher de Kaneohe, no Havaí, foi indiciada por supostamente tentar desviar dinheiro da empresa de segurança na qual trabalhava, registrando e recebendo o pagamento de dois funcionários fictícios. O golpe teria rendido mais de US$ 200 mil. Um segundo método é a fraude tipo salame, que tem esse nome porque envolve tirar pequenas fatias de cada vez. No caso, são retiradas frações nos cálculos de arredondamento ou nas transações de troco, desviando-se alguns centavos. Quando a empresa lida com milhões de transações, isso pode se transformar em uma fraude considerável. Uma terceira forma de fraude é o pagamento de bens inexistentes. Isso pode acontecer quando uma pessoa tem a capacidade de fazer um pedido de compra, receber mercadorias e emitir cheques. As pessoas têm um problema parecido quando um vendedor na internet recebe dinheiro antecipado por produtos que não existem ou quando participam de leilões fraudados por meio de lances falsos para aumentar os preços. À medida que aumentou o uso de sistemas de computação, o mesmo ocorreu com as ameaças. Surgiram formas mais sofisticadas de fraude, muitas vezes aproveitando o enfraquecimento dos controles no ambiente de TI. Nos primórdios da computação, jovens se deixaram fascinar pelo desafio de usar um computador e um modem para invadir outros computadores, e assim começou a era dos hackers adolescentes. Era mais pelo desafio individual e reconhecimento dos colegas, embora houvesse os primeiros exemplos do que hoje chamamos de ciberespionagem e cibercrimes. Um hacker famoso da década de 1990 foi Kevin Mitnick, também conhecido como Condor. Depois de uma década de ataques apenas para demonstrar suas habilidades, Mitnick finalmente foi pego e condenado a três anos de prisão. A história completa, com livro e filme, está descrita no site Takedown. À medida que o uso da web aumentou e os sites de informação e portais de negócios se tornaram mais comuns, a desfiguração de sites para fins políticos se tornou uma nova ameaça à segurança. Nas primeiras desfigurações, a mensagem dos hackers era do estilo “te peguei”. Mais recentemente, porém, as desfigurações têm sido usadas para transmitir uma mensagem política. À medida que foi crescendo a dependência dos sistemas de TI, outra ameaça à segurança que surgiu foi o ataque de negação de serviço, em que um invasor remoto compromete os sistemas de TI explorando vulnerabilidades ou sobrecarregando sua capacidade de lidar com o volume dos fluxos de informações. Com o crescimento da internet, os adversários conseguiram assumir o controle de um grande número de computadores, chamados de botnet, e concentrá-los em um único alvo, ampliando consideravelmente o impacto da negação de serviço. Isso é conhecido como ataque distribuído de negação de serviço, ou DDoS. Um exemplo mundial de DDoS ocorreu em fevereiro de 2014, quando um invasor desconhecido lançou uma onda contínua de ataques distribuídos de negação de serviço em vários alvos, país por país, no mundo todo. A escala do ataque foi imensa. O ataque usou um recurso especial do protocolo NTP para amplificar os dados. Quando os pacotes atingiram o alvo, eles foram amplificados 50 vezes, resultando no equivalente a cerca de 250 mil ataques de negação de serviço individuais. Uma das empresas-alvo dessa campanha foi atacada por 4.278 endereços IP individuais de mais de 80 países, distribuindo um fluxo contínuo de mais de um milhão de pacotes por minuto durante cerca de uma hora. O gráfico mostra o cronograma dos dados chegando da internet a esse alvo.

**Camuflagem e fluxos de dados alternativos**

As pessoas que invadem sistemas com malwares fazem de tudo para ocultá-los quando atingem o alvo. Se elas conseguirem evitar a detecção, é mais provável que alcancem seus objetivos. Vamos ver algumas maneiras de ocultar malwares. O primeiro método são as técnicas usadas pelo sistema operacional Windows para ocultar suas atividades. Um exemplo é o arquivo oculto usado pelo Internet Explorer para manter uma cópia do histórico de navegação do usuário. Cada versão do Explorer oculta o arquivo em um lugar diferente. Veja onde fica o meu. Estou em um shell de comando. Vou entrar no diretório de aplicativos do usuário digitando cd \users\user\appdata\local e navegar até o diretório microsoft\windows. Ao listar o conteúdo do diretório, existem vários arquivos e pastas, mas não há um subdiretório History. Mesmo se eu incluir arquivos ocultos, não vemos um diretório History, mas o subdiretório existe – só que não podemos vê-lo. Vamos tentar mudar de diretório e acessá-lo. Deu certo. Agora vamos ver o que temos aqui. Vemos o arquivo desktop.ini. É o método que a Microsoft usa para ocultar o subdiretório. Agora vamos ver o que tem dentro dele. Existem duas entradas de camuflagem. A primeira é uma linha CLSID que impede que o subdiretório seja incluído em buscas baseadas em arquivos, e a segunda, a linha UICLSID, impede que o subdiretório seja visto usando o Windows Explorer. Outra maneira pouco conhecida de ocultar informações no disco é usar o chamado fluxo de dados alternativo. Nos primeiros sistemas de arquivos MS-DOS e FAT, os arquivos eram simplesmente sequências de dados que podiam ser lidas byte por byte pelos aplicativos. No NTFS, um arquivo é uma estrutura complexa. Os arquivos NTFS contêm, no mínimo, uma seção chamada $DATA, onde ficam os dados lidos por um aplicativo. Esse é o fluxo de dados. Porém, um arquivo pode ter muitas outras seções, cada uma com seu próprio nome e capaz de armazenar informações. São os fluxos de dados alternativos. Convém destacar que o Windows reconhece apenas a seção $DATA, portanto os dados em qualquer fluxo de dados alternativo não são reconhecidos. Voltando ao shell de comando, criei um novo arquivo chamado datafile.txt. Vamos ver como fica ao digitá-lo. É um arquivo de texto simples. Vamos conferir o tamanho dele. Tem 96 bytes. Também criei outro arquivo chamado adsfile.txt, que é minha mensagem secreta. Usando a técnica de fluxos de dados alternativos, posso inserir isso em um fluxo de dados oculto em datafile.txt digitando type adsfile.txt para datafile.txt:hidden.txt. Agora vamos ver como ficou o datafile. Aparentemente, nenhuma alteração. O tamanho também não mudou. Porém, se eu digitar more datafile.txt:hidden.txt, veremos o texto oculto. Os fluxos de dados alternativos podem ser usados para ocultar arquivos executáveis. Como exemplo, vou inserir a calculadora do Windows neste arquivo de texto. Vamos verificar o datafile.txt de novo. Veja que não há nenhuma alteração. Podemos usar o comando wmic do Windows para executar isto. Observe na lista de tarefas que a calculadora está em execução. Embora os fluxos de dados alternativos não possam ser vistos na interface normal do Windows, a linha de comando permite usar a opção /r no comando dir para visualizá-los. Agora vemos os dois arquivos adicionais e os fluxos de dados alternativos hidden.txt e mycalc.exe.

Arquivo oculto no Windows.

**C:\Users\Suporte Técnico\AppData\Local\Microsoft\Windows\History>type desktop.ini**

**[.ShellClassInfo]**

**ConfirmFileOp=0**

**CLSID={FF393560-C2A7-11CF-BFF4-444553540000}**

**UICLSID={7BD29E00-76C1-11CF-9DD0-00A0C9034933}**

**Como ocultar o uso de processos**

Alguns malwares se escondem bem na nossa frente. Veja um malware simples que usa uma técnica conhecida como ganchos do Windows para interceptar pressionamentos de teclas. No Windows, quando uma tecla é pressionada, um evento é transmitido para o sistema operacional. O Windows usa o driver do teclado para ler o caractere que foi pressionado e o envia como uma mensagem para o aplicativo que está aguardando. No entanto, o Windows também permite que outros processos analisem a mensagem à medida que atravessa o sistema. É assim que funcionam as teclas de atalho. O Windows faz isso usando algo chamado rotina de retorno de chamada do tipo gancho de teclado. Ela captura a tecla que foi pressionada e, então, a chama de volta após processá-la, repassando-a ao seu destino inicial. Usando essa técnica, bastam algumas linhas de código para escrever um simples registrador de pressionamentos de tecla. Eu preparei um programa com ganchos de teclado chamado intercept.cpp, que vemos aqui no WordPad. É um programa bem simples. As cinco primeiras linhas representam o código de configuração padrão de um programa de C++. Em seguida vem uma linha de código que é executada quando o programa é iniciado. Ela abre um arquivo de saída chamado incept.text, que vamos usar mais tarde para armazenar caracteres interceptados. As seis linhas seguintes especificam um procedimento de retorno de chamada. É o código executado quando uma tecla é pressionada. Os eventos no Windows são bastante granulares. Por exemplo: um pressionamento de tecla envolve dois eventos – KeyDown e KeyUp, ou seja, pressionar a tecla e soltar a tecla. Nós só precisamos verificar um deles, então a rotina de retorno de chamada verifica apenas o evento WM\_KEYUP. Ao ver isso, ela grava no arquivo de saída os dados associados a esse evento, que é o código interno do Windows para o caractere pressionado no teclado. A última ação dessa rotina é repassar o evento, chamando o próximo gancho da cadeia com a função CallNetHookEx, que tem um nome bastante adequado e permite que a mensagem retome seu caminho até o aplicativo de destino. O restante do programa é a lógica de controle principal do processo do Windows que executa a rotina de retorno de chamada. Convém fazer algumas observações. Podemos ver a chamada SetWindowsHookEx, para se conectar ao teclado. Também observamos um atalho sendo registrado, Alt + 9. Abaixo disso, o código faz um loop, esperando que o atalho Alt + 9 seja pressionado, e nesse momento ele é desconectado, fecha o arquivo de interceptação e é encerrado. Certo. Já compilei o código, então vamos ver o programa em ação. Agora vamos fechar o shell de comando. Estou com meu portal bancário pronto, então vamos digitar as credenciais. Vamos fazer uma análise forense simples para ver o que está sendo executado no computador. Vou digitar Control + Alt + Del para abrir o gerenciador de tarefas. Não existe nada na lista de aplicativos que mostre que o interceptor está sendo executado, mas vamos abri-la para ver mais detalhes. Ainda não está aparecendo nos aplicativos, mas, se eu rolar mais para baixo, vejo o intercept.exe. Está bem à vista para quem procurar, mas, com um nome não descritivo, ninguém notaria. Pressionando Alt + 9, vemos que o processo intercept é encerrado. Voltando ao shell de comando, vamos dar uma olhada no arquivo de registros que o programa intercept está usando. Aqui vemos o que o programa interceptou. Essa interceptação contém outras atividades do teclado, como as teclas Alt e Shift, exibidas como vários caracteres especiais. Mas nosso invasor registrou o que desejava – as credenciais da nossa conta bancária. Os ganchos de teclado são apenas um dos muitos métodos de captura com ganchos ou anzóis. A lição aqui é que, quando a vítima morde o anzol, seu sistema e possivelmente todas as suas senhas e códigos de acesso online ficam comprometidos.

**Controle do alvo por meio de um rootkit**

Os malwares que usam as técnicas básicas de ocultação e camuflagem podem ser detectados por um investigador capacitado. Consequentemente, os invasores mais sofisticados desenvolveram técnicas em que malwares robustos, não apenas como um aplicativo ou processo, analisam o sistema operacional a fundo. Esse tipo de malware é chamado de rootkit. Para implementar um rootkit, primeiro um invasor precisa penetrar no sistema de um alvo e, em seguida, usar o chamado dropper para instalar o rootkit, que ele carrega como payload ou baixa posteriormente. A função do dropper é verificar se o rootkit já existe no sistema, se o sistema está operando dentro de uma máquina virtual e fazer verificações especiais, como o país em que está operando. Então, depois de confirmar que o alvo é legítimo e está disponível, ele insere o rootkit no sistema, assegura que possa ser reiniciado após a inicialização do sistema e inicia sua execução. Um rootkit não explora vulnerabilidades. Ele é feito para ocultar, operar e executar sua missão usando funções normais do sistema. Da mesma forma, um rootkit não é um vírus – é um implante que existe em um alvo, mas não se propaga sozinho. No entanto, ele pode ser combinado com um código semelhante a um vírus para permitir a propagação lateral depois de instalado em um sistema. Um rootkit é projetado para contornar sistemas de detecção de intrusão. Por exemplo: ele pode conter um código para procurar e desabilitar certas formas de antivírus ou softwares de detecção de intrusão baseados em host. Ele também precisa evitar a detecção por analistas forenses, e o melhor lugar para se esconder é no kernel do sistema operacional. Trata-se do núcleo interno do sistema operacional, e os sistemas de detecção de intrusão não conseguem examinar a parte interna do kernel com facilidade. Para entrar no kernel, os rootkits precisam ser codificados como uma forma especial de programa chamada de módulo ou driver carregável do kernel. A Microsoft fornece um kit de desenvolvimento de drivers para Windows, ou DDK, que permite desenvolver módulos carregáveis do kernel. Um driver usa técnicas de programação bastante complexas, então não vou me aprofundar no processo de desenvolvimento e implementação de drivers. Apenas digo que os drivers operam na parte mais profunda do kernel, chamada anel zero, e isso lhes dá acesso a todas as estruturas de dados do kernel. Além disso, por estar no kernel, os rootkits são executados com privilégios elevados. Há muito pouco que um rootkit não consiga fazer, e pode ser bem difícil encontrá-lo. Por ser um processo privilegiado no kernel, um rootkit pode fazer a manipulação direta de objetos do kernel, ou DKOM. O kernel usa estruturas de dados para monitorar o ambiente. Um exemplo é a estrutura de dados de processos mantida no módulo de processos. Trata-se de uma lista duplamente vinculada, o que significa que cada entrada tem dois ponteiros. A cadeia de ponteiros para a frente começa no topo da lista, e cada entrada aponta para a entrada seguinte da lista. A cadeia de ponteiros para trás começa no final da lista, e cada entrada aponta para a anterior. Usando esses ponteiros, o kernel consegue monitorar e gerenciar todos os processos ativos. Quando um usuário abre o gerenciador de tarefas, o que realmente acontece é que o aplicativo do gerenciador de tarefas é carregado e, então, chama o kernel solicitando uma lista de aplicativos e processos. O kernel verifica suas estruturas de dados, cria as listas e as envia de volta ao gerenciador de tarefas para exibição. Quando o rootkit é carregado, uma das primeiras coisas que faz é examinar a lista de processos. Ele segue a cadeia para a frente para encontrar seu próprio processo, o anterior e o seguinte. Depois de obtê-los, ele pode alterar os ponteiros, de modo que o processo anterior aponte para o seguinte ao rootkit e o processo seguinte tenha seu ponteiro para trás alterado e aponte para o anterior ao rootkit. O primeiro rootkit conhecido por realizar a DKOM surgiu por volta de 2006 e fez exatamente isso – alterou os ponteiros da lista nas estruturas de dados de processos para contornar a própria entrada. Com isso, as atividades do rootkit ficam ocultas no gerenciador de tarefas e no agendador de eventos. A DKOM pode ser usada nas estruturas de dados, como a lista de drivers e a lista de portas abertas. Os rootkits também interceptam e removem suas entradas nas listas de diretórios e arquivos. Uma maneira comum de permitir comando e controle remotos é criar um canal criptografado por meio do uso de conexões Secure Shell. Isso dá ao invasor controle remoto sobre sistemas comprometidos e, ao mesmo tempo, criptografa qualquer malware que possa ser baixado, evitando que seja detectado por sistemas de detecção de intrusão baseados em rede, uma ferramenta de monitoramento. O uso do SSH tem a vantagem de exigir a entrada de um nome de usuário e senha, garantindo assim que apenas o adversário possa usar o backdoor para acessar o rootkit. O Trojan Downloader 3, ou TDL3, é a terceira geração de rootkit desenvolvida pelo grupo de cibercrime Dogma Millions e é um exemplo de rootkit do mundo real. O rootkit se adiciona como um driver de impressora, conferindo-lhe privilégios de driver de modo kernel. Ele instala um sistema de arquivos criptografado para uso próprio que começa no final do disco rígido e se expande para trás até o início do disco. O Windows considera que isso é apenas espaço livre, vazio, o que significa que não é detectado pelas técnicas tradicionais de varredura. O TDL3 opera em um modelo de custo por instalação, registrando quando é instalado em um alvo e quem adquiriu a versão de origem. O rootkit TDL3 é usado para baixar, instalar e ocultar módulos de payload fraudulentos que podem então fazer o monitoramento dos pressionamentos de tecla, realizar ataques disruptivos de negação de serviço e muitas outras ações.