Alta Disponibilidade: Confiabilidade e Alta Disponibilidade: Principais Desafios

Resumo

Referência bibliográfica

- SOARES, Juliane Adélia... [et al.]. **Redes de alta disponibilidade**. – Porto Alegre: SAGAH, 2020.

As Situações de Falhas de Rede

Redes de computadores, como toda e qualquer tecnologia, são compostas de diversos componentes que possuem complexas interações e diversos níveis de confiabilidade. Somada a própria tecnologia, as redes possuem em sua configuração, protocolos e usuários como fontes de falhas e indisponibilidade.

Desta forma, para que uma rede funcione em sua plena capacidade, deve começar com seu hardware, ou seja, componentes de qualidade, com elevada confiabilidade e capacidade de reparo e recuperação rápidos. Portanto podemos classificar o hardware como a primeira fonte de indisponibilidade de uma rede, e de acordo com Schmidt (2006; p. 44),

> A primeira falha a ser considerada ao se estudar a disponibilidade de rede decorre de problemas na camada física, visto que todas as demais camadas deverão, inevitavelmente, depender desta, de modo que ela sofrerá maior impacto de falhas. Tanto o hardware (dispositivos) como o software (firmware, software embarcado, etc.) envolvidos nos equipamentos que compõem a camada física podem ser responsáveis por alguma falha ou indisponibilidade do serviço.

Desta forma, devemos compreender que se os elementos de ligação entre os dispositivos falham, sua conexão é interrompida, o que coloca boa relevância em componentes como o cabeamento da rede, seus modens, roteadores e switches. Assim, quando se necessita criar redundância, tais componentes podem ser facilmente duplicados, ao contrário de servidores que possuem custo elevado.

O cabeamento de rede costuma causar inúmeros problemas de conflitos e indisponibilidade da rede pois é um elemento facilmente negligenciado, portanto são frequentes os problemas de rompimento de cabo e até mesmo mal contato nos seus conectores, denominados RJ-45. Conexões de cabos de rede podem ser facilmente duplicadas, pois mesmo que não sejam imediatamente conectadas ao dispositivo, ficam à espera de um problema localizado para entrarem em ação. A falta de conexão devido a problemas de cabeamento causa indisponibilidade geral ao sistema, conforme apresentam Marcus e Stern (2003; p.45),

> A falta de conexão tem o agravante de ocorrer de maneira silenciosa, ao contrário das falhas por sobrecarga. Sem a conexão, é impossível retornar até mesmo uma mensagem de erro. Nesse caso, em vez de apenas testar o endereço final, pode-se recorrer a um comando muito utilizado tanto no sistema Windows quanto no Linux, o netstat, que fará a contagem de pacotes para auxiliar na determinação do local da falha.

Além do cabeamento, outros problemas podem ter sua origem na parte física das redes: as falhas em componentes internos a dispositivos embarcados, falha e erros de configuração, sobrecarga na manipulação dos pacotes, dentre outras. O que deve ser evitado a todo custo são as falhas ocasionadas por componentes de baixa qualidade que oferecem baixa confiabilidade, um parâmetro vital para a alta disponibilidade.

Hardware de rede de baixa confiabilidade pode facilmente causar interrupções a rede por gerar elevadas taxas de latência, tempo de resposta devido a sua má gestão do tráfego de pacotes nas redes. Agora busque pensar também, em redes maiores, como a internet, por exemplo, a maior rede de todas, desta forma, pense como estes problemas podem impactar no seu uso em geral? Desta forma, o efeito da indisponibilidade da internet é apresentado por Marcus e Stern (2003; p. 46),

> Na internet, que conecta diferentes redes locais a uma rede maior, enquanto as redundâncias mais próximas dos destinatários tendem a ampliar o custo, longe desses pontos, a redundância é atingida praticamente de maneira natural através das múltiplas rotas possíveis nela. Um dos problemas enfrentados em roteamento é o surgimento de rotas assimétricas, em que, apesar de o cliente conseguir enviar um pacote para um servidor, este desconhece a rota para retornar uma resposta ao cliente. As mudanças de rotas envolvem processamento e, certamente, não incorrerão em rotas de mesma velocidade, mas é sempre preferível utilizar uma rota mais lenta que esteja disponível do que permanecer incomunicável.

O cabeamento redundante ajuda a eliminar falhas mais triviais, mas em uma era de conectividade sem fios, é preciso que a confiabilidade da rede seja aprimorada com o uso de roteadores repetidores de sinal. E vale ressaltar que embora componentes de maior confiabilidade tenham custo maior, em situações e sistemas que dependem da alta disponibilidade por questões de segurança e continuidade.

2. Principais Ameaças a Alta Disponibilidade

Seja a natureza, a infraestrutura predial, o software, sua configuração, o hardware e sua instalação e até mesmo o fator humano com seu uso e o tratamento que dá a segurança, são inúmeros os elementos que potencialmente podem causar danos a uma rede e a tornar indisponível. Mas para cada fator existem as formas de se detectar a falha que origina, o que torna o processo de monitoramento e prevenção de falhas complexo e custoso. De acordo com Critchley (2014, p. 14), temos nas falhas de comunicação um dos aspectos mais difíceis de se detectar a origem:

> As causas de uma falha de comunicação na rede são difíceis de se determinar, uma vez que existem muitos fatores envolvidos, bem como o fracionamento das mensagens, mudanças de rota, múltiplas rotas e outros fatores envolvidos. No entanto, é comum que os protocolos de comunicação sejam escolhidos de acordo com o menor risco envolvido, visto que o que pode ser uma garantia de menor disponibilidade para uma determinada aplicação pode não ser para outra.

A parte lógica da rede é feita através dos protocolos de comunicação que apresentam diferentes características e desta forma, resolvem determinados problemas, pois para as métricas de taxa de transmissão e a garantia de entrega não estão presentes em um mesmo protocolo. Neste caso temos o protocolo UDP (User Datagram Protocol - Protocolo de Datagrama de Usuário), capaz de priorizar a velocidade de transmissão, ao passo que o protocolo TCP garante a entrega com o processo de confirmação de recebimento e de acordo com Critchley (2014, p. 49) temos,

> À primeira vista, pode parecer que o TCP sempre será vantajoso perante o UDP. No entanto, em serviços de streaming, ou qualquer outro serviço que necessite trocar informações em tempo real, os pacotes perdidos são menos importantes que o atraso da informação. Por exemplo, você dificilmente conseguiria participar de um jogo virtual, interagindo com outros jogadores, caso buscasse garantir que nenhuma informação fosse perdida. (CRITCHLEY, 2014; p.49).

Mesmo que sejam escolhidos os melhores protocolos, ainda existirão outros elementos a serem aprimorados para que a indisponibilidade seja evitada. Em geral a indisponibilidade é ocasionada por mais de um fator em conjunto, como podemos compreender na lista apresentada a seguir:

1. Falhas parciais ou completas de algum componente de hardware/ software da rede;
2. Quedas de energia elétrica em algum ponto da rede;
3. Falhas no sistema embarcado de algum equipamento;
4. Manutenção agendada inevitável;
5. Erros de configuração;
6. Desastres naturais. (CRITCHLEY, 2014: p.50)

Desta forma, o processo de se mitigar as falhas de rede invariavelmente será multidisciplinar, demandando ações e processos claros para os diversos aspectos da rede. Manter um cabeamento de qualidade não evita todos problemas, pois em casos onde o acesso sem fio for majoritário, a estrutura de cabos físicos não terá grande efeito. Assim como a utilização de mais repetidores pode melhorar o acesso de dispositivos que demandam a conexão sem fio, ao passo que não ajuda no desempenho geral da rede se ocorrer o rompimento do cabo que leva a internet ao roteador, modem.

3. Minimizando as Ocorrências de Indisponibilidade

Fazendo um paralelo, veículos somente são classificados como confiáveis se seus milhares de componentes forem igualmente confiáveis. Em uma competição de Rally, onde veículos cruzam em pistas de piso irregular, vários componentes são exigidos no limite de suas especificações, portanto são componentes desenvolvidos para suportar tais condições.

Desta forma, se uma rede é projetada para ser altamente disponível, deve conter componentes, configurações e processos que suportem estas especificações. Desta forma, atacar a indisponibilidade está relacionado a elevação das características dos elementos que fazem parte de uma rede. Claro que afirmar que melhores componentes fazem uma rede melhor, não induz a compreensão das ações necessárias.

Uma forma de se começar a agir na mitigação da indisponibilidade está na busca pela eliminação de erros e problemas da rede. Esta varredura ajuda a detectar problemas e até mesmo deficiências, casos de hardware subdimensionado ou de recursos mal distribuídos. Portanto, antes de se investir em bons equipamentos, deve-se corrigir os problemas da rede. O próximo passo está focado na aquisição de componentes que agreguem robustez e desempenho constante, sólido a rede, e sua escolha é complexa,

> Não apenas a confiabilidade deve ser considerada na hora de selecionar os dispositivos que irão compor a rede, como também a reparabilidade destes, já que mesmo manutenções agendadas não podem ser evitadas e, certamente, impactam de alguma forma o usuário. Dependendo das técnicas e dos métodos utilizados, a reparabilidade é capaz de contornar falhas difíceis de serem controladas. Por exemplo, apesar de não ser praticável, na maior parte dos casos, o uso de redundâncias em componentes internos de processamento é completamente possível, bem como a utilização de equipamentos redundantes, alocados de forma a facilitar a troca imediata de um equipamento danificado por outro funcional, em vez de se tentar repará-lo antes de restaurar o serviço. (SOARES ET AL. 2020, p. 51).

Para provedores de serviços na nuvem, como hospedagem de sites, drives virtuais, existe um documento que orienta o nível de robustez necessário, as ações e investimentos que se fazem urgentes para que seja cumprido, estamos tratando do acordo de nível de serviço, conhecido pela sua sigla SLA (*Service Level Agreement*).

Este documento é importante para que o provedor mantenha em foco o que deve promover em termos de monitoramento de desempenho da rede e investimento em infraestrutura para manter o serviço adequado e conforme promete em seu contrato. O SLA é usado pelo cliente para saber o que esperar sobre a performance e disponibilidade do serviço que contrata e da parte do provedor, como compromisso e mecanismo de justificativa para processos e investimentos em alta disponibilidade.

O SLA é uma espécie de medida de performance, como um indicativo para que as ferramentas de monitoramento tenham um parâmetro a seguir, conforme Soares et al. (2020, p. 50) complementa:

> As ferramentas de controle e monitoramento podem auxiliar, inclusive, na antecipação de falhas. Por exemplo, conhecendo-se o perfil sazonal de transmissão de dados, é possível prever que um aumento recente no consumo poderá resultar em um volume de dados maior do que a capacidade de entrega da rede em outro período, antecipando, assim, a intervenção, que poderá ser realizada de maneira controlada e com menor impacto. (SOARES ET AL. 2020, p.52).

Vale ressaltar que a indisponibilidade pode ser ampliada por elementos que não são necessariamente problemas ou erros, como a configuração de redes privadas ou VPN´s, pois sua existência promove a restrição do acesso e limita a existência de rotas alternativas, algo ruim no caso de falhas na rede principal. O uso de VPN´s é comum em clientes de provedores de serviços em nuvem e pode ser descrito como,

> Em uma VPN, os dados utilizam a mesma infraestrutura comum à internet, mas somente são compreendidos pelos membros dessa rede. Assim, por mais que os dados estejam acessíveis a outros dispositivos ou aplicativos, em virtude da criptografia utilizada, somente os membros que possuírem as chaves para decodificar os dados poderão interpretá-los adequadamente. O tunelamento utilizado consiste em encapsular o pacote original de dados em outro pacote, cujo destino é o membro da rede que deverá receber a mensagem criptografada. (SOARES ET AL. 2020, p.52).

A indisponibilidade deve ser tratada como um indicador de que alguns dos componentes, elementos de uma rede apresentam falhas, subdimensionamento ou configuração equivocada. Para os provedores de serviços web a indisponibilidade, quando baixa, é um indicador de qualidade.

4. Importância da Alta Disponibilidade nas Redes

Embora a indisponibilidade represente um problema sistêmico, ou seja, afeta todos usuários de uma rede, a alta disponibilidade nem sempre é percebida, ao menos pela perspectiva do usuário. Compreendendo que alta disponibilidade é muito mais do que um serviço estar constantemente disponível e em sua qualidade nominal, não estamos falando em desempenho, ou seja, uma rede altamente disponível não oferece necessariamente, velocidade de transmissão superior às demais redes.

Mas à medida que mais serviços são criados em ambientes web e outros tantos migram para ele, começamos a compreender que alta disponibilidade representa a sustentabilidade destes serviços, sem ela é possível que muitas das funcionalidades e serviços da internet não tivessem a grande procura que tem hoje.

A esmagadora maioria das empresas usa internet, redes internas em algum momento dentro de seus processos, e esta presença digital aumenta a cada dia, desta forma a alta disponibilidade começa a ser vista como uma dependência nestas situações e assim, temos diferentes níveis de criticidade a momentos onde a empresa fica sem sua conexão com a internet,

> Em alguns casos, em que a produção e a atividade são locais e internas ao ambiente da organização, é possível que algumas aplicações funcionem bem durante períodos sem conexão. No entanto, essa não é a situação na maioria dos casos, pois uma simples ordem de serviço ou monitoramento de uma produção não sobreviverá à falta de conexão por tempos prolongados, já que as mudanças na produção, apesar de serem lentas, ocorrem e demandam ajustes para que uma maior produtividade possa ser obtida. (SOARES ET AL. 2020, p.53).

Estudante, já parou para refletir sobre a importância das redes, da computação nas empresas? Quais estabelecimentos comerciais você conhece que não usam nenhuma forma de rede ou computador? Em empresas de pequeno e médio porte já podemos começar a perceber as redes como um recurso vital para a execução de suas atividades. Portanto, podemos afirmar que as empresas dependem de alta disponibilidade para ter sucesso em suas atividades? Para Soares et al. (2020, p.54), as redes promovem uma série de benefícios empresariais,

> As redes possibilitam ampliar o poder de processamento e acessar serviços indisponíveis de outra maneira de forma rápida. Quanto mais integradas as máquinas estão, melhor tende a ser o aproveitamento do processamento disponível, já que, nesse caso, o usuário pode solicitar o uso de alta capacidade computacional em um breve instante de tempo, sem

necessariamente recorrer à expansão da sua própria infraestrutura, o que demandaria custos e investimentos elevados.

São incontáveis os relatos de pessoas que afirmam trabalhar ou terem trabalhado em empresas cujos problemas de rede eram frequentes, diários e severos ao ponto de impedir o andamento das diversas funções que executam. Com o crescimento da adoção de recursos e serviços na nuvem a conexão de rede passa a ser um item ainda mais crítico para a empresa, ao passo que sua indisponibilidade pode a incapacitar totalmente.

Exercícios

1. Os modelos de arquitetura em camadas de uma rede de computadores, foram criados como forma de reduzir a complexidade dos processos que envolvem os inúmeros aspectos de hardware, softwares, controle de erros, integração como sistema operacional, entre muitos outros. Este modelo consiste dividir o projeto de rede em funções independentes, de forma que estas sejam agrupadas em camadas, considerando os modelos OSI (*Open Systems Interconnection*) e TCP/IP ((Transmission Control Protocol/Internet Protocol), ambos modelos apresentam sua distribuição em camadas, apresentando como sua primeira camada, a física.
Neste caso, qual a importância de se considerar primeiramente a camada física, na análise das falhas na disponibilidade de uma rede?

a) Pois a camada física compreende os softwares da rede, que são responsáveis diretos pela disponibilidade.
b) Pois as demais camadas dependem, inevitavelmente, da integridade da camada física para manter a disponibilidade.
c) Pois a camada física depende diretamente das demais camadas, apesar de sofrer menos impactos dentre todas as camadas.
d) Pois os softwares que compõem a camada física são responsáveis por conectar fisicamente as máquinas que integram esta rede.
e) Porque a camada física é a única camada que suas falhas podem resultar em indisponibilidade computacional.

2. Em uma rede de computadores, as falhas de conexão que causam indisponibilidade computacional podem ser de diferentes origens, entre elas falhas na camada física da rede. Esta falha pode ocorrer em diferentes dispositivos, mas em casos de redes internas, também podem ter sua origem em problema no cabeamento, e normalmente ocorrem de maneira silenciosa.
Considerando a indisponibilidade computacional como resultado de falhas de cabeamento em redes internas, como deve ser o cabeamento, de modo que esta indisponibilidade seja mitigada?

a) O cabeamento não é mais utilizado, deve ser substituído, componentes sem fio.
b) Deve ser redundante, ou seja, deve possuir rotas de acesso alternativas.
c) Devem possuir cabeamento único, ou seja, único caminho.
d) Deve possuir uma estrutura simples e concisa.
e) Devem estar conectados fisicamente com repetidores de sinais.

3. Para os usuários das redes de computadores, a disponibilidade é um elemento de absoluta importância, uma vez que na ocorrência de falhas nesta propriedade, as redes perdem sua funcionalidade primordial, que é representada pela transmissão efetiva de dados. É comum que os protocolos de comunicação sejam escolhidos de forma a minimizar os riscos de problemas e falhas, de acordo com o que se espera dos serviços de uma rede. Os protocolos TCP (*Transmission Control Protocol*) e UDP (*User Datagran Protocol)* determinam de que modo os dados serão trocados na internet, cada uma com suas funcionalidades, e apesar das aparentes vantagens na utilização do protocolo TCP, isso devido também a sua popularidade, o UDP é de grande valia em determinadas situações.
Em qual situação o protocolo UDP apresenta mais vantagens que o protocolo TCP?

a) Em serviços que necessitem de garantia dos pacotes de entregas.
b) Quando a entrega exige uma efetiva confirmação
c) Em serviços que necessitem de troca de informações em tempo real.
d) Em serviços que exijam exatidão de dados.
e) Quando são necessários o reenvio de dados.

4. A disponibilidade de uma rede de computadores representa seu foco principal, sem o qual esta rede tenha qualquer funcionalidade, uma vez que a ausência deste quesito básico impede que a transmissão de dados seja possível. Existem diversas medidas que podem ser tomadas para minimizar a ocorrência de indisponibilidade na rede, dentre elas podemos destacar a eliminação de falhas localizadas que são resultado, normalmente de equipamentos de má qualidade, e dessa forma promover a confiabilidade. Entretanto, não é prudente focar apenas na confiabilidade dos dispositivos de uma rede, é necessário considerar sua reparabilidade, e na implementação das chamadas SLA (*service level agrément*).
Assinale a alternativa que apresente o que representam os SLAs?

a) Acordos apresentados ao prestador de serviços, em que constam todas as demandas dos usuários para dimensionamento dos serviços oferecidos.
b) Contrato de prestação de serviços em que são estabelecidos os acordos de consumo e pagamento de acordo com o perfil do usuário.
c) São acordos de nível de serviço, que apresentam ao usuário o entendimento sobre indisponibilidade, além de todas as informações relacionadas ao serviço prestado.
d) Contrato de contratação de serviços em que o usuário o tempo em que a rede será utilizada, assim o prestador de serviços pode realizar a programação de serviços.
e) São acordos estabelecidos entre os prestadores de serviços e os órgãos governamentais, que envolvem a LGPD (Lei Geral de Proteção de Dados).

5. A disponibilidade computacional de um modo geral deve ser entendida como uma característica de um sistema computacional, e representa um conceito que vai além de um produto ou uma aplicação que pode ser facilmente instalada e executada. E quanto a alta disponibilidade, adentramos um assunto ainda mais complexo, pois esta característica ainda apresenta caráter resistente a falhas. Dessa forma, redes de alta disponibilidade dispõe de diversos tipos de ferramentas que viabilizam esta solução, uma delas são as ferramentas de controle e monitoramento da rede.
Qual a importância das ferramentas de controle e monitoramento da rede para a manutenção da alta disponibilidade?

a) Estas ferramentas são capazes de antecipar falhas, detectando sua origem de forma rápida, evitando desta forma uma possível indisponibilidade.
b) As ferramentas de controle e monitoramento são capazes de realizar correções automáticas no sistema, por meio de inteligência artificial.
c) Estas ferramentas promovem a reparabilidade do sistema através de dispositivos e componentes adequados.
d) Ferramentas que são capazes de eliminar toda e qualquer falha localizada no sistema, que pode resultar em indisponibilidade.
e) Promovem a criptografia de dados, diminuindo assim a possibilidade de falhas que possam causar indisponibilidade.

6. Os acordos de nível de serviços são de grande funcionalidade para o relacionamento dos usuários com os prestadores de serviços para que fique entendido os termos de serviços indisponíveis. Estes acordos, ou (SLA, Service level agreement), são importantes pra estabelecimento de um relacionamento claro entre o usuário e o prestador de serviços, inclusive ao que diz respeito às garantias de privacidade e sigilo das informações, que podem receber um incremento através da criação de uma VPN (Virtual Private Network; ou Rede Privada Virtual).
Assinale a alternativa que apresente a funcionalidade das VPNs (Virtual Private Network)?

a) São redes caracterizadas pela alta disponibilidade, composta por dispositivos de alta tecnologia e que são funcionais para comunicação empresarial.
b) São redes virtuais, criadas dentro da infraestrutura de uma rede, que promovem a criptografia dos dados, promovendo a privacidade e o sigilo dos dados.
c) São redes privadas que têm como maior objetivo a privacidade dos dados, e para possuem estrutura própria, não relacionada com uma infraestrutura de rede padrão.
d) Redes privadas que estão interligadas por cabeamento interno, sem acesso a internet de forma a manter a privacidade dos dados.

Gabarito

1. Letra B.
   A alternativa está correta pois todas as camadas que compõem a arquitetura de uma rede dependem da integridade da camada física, pois uma falha em um equipamento ou mesmo um simples rompimento de um cabo pode causar indisponibilidade.

2. Letra B.
   A alternativa está correta pois em redes internas que não se utilizam de dispositivos sem fio, é necessário que o cabeamento seja redundante, ou seja que deve oferecer rotas de acesso alternativas, sendo pouco provável que a rede ocupe toda estrutura quando ocorrer algum tipo de falha.

3. Letra C.
   O protocolo UDP é mais vantajoso que o TCP em casos em que a taxa de transmissão é mais importante que a garantia de entrega dos pacotes. Um exemplo é em casos de serviços de streaming ou qualquer outro serviço que necessite a troca de informações em tempo real.

4. Letra C.
   Alternativa está correta pois o SLA (*service level agrément)* representa os acordos de nível de serviços que deve ser implementado de forma que fique claro o que deve ser entregue pelo prestador de serviços ao usuário, neste acordo deve estar contido o que ser entendido por indisponibilidade, além das demandas de segurança, momentos de maior impacto, entre outras informações pertinentes a prestação de serviços.

5. Letra A.
   A alternativa está correta pois os sistemas de controle e monitoramento de rede são capazes de antecipar falhas, sua origem e causa de forma rápida, o que terá grande impacto no tempo de resposta, ou seja, promovendo a restauração antes que cause algum tipo de problema que leve a indisponibilidade.

6. Letra B.
   A redes VPN (Virtual Private Network) são redes virtuais, que se utilizam da infraestrutura de uma rede, e tem a funcionalidade de criptografar dados por meio técnicas de tunelamento, e que apesar de utilizar uma infraestrutura de rede comum, somente os membros desta VNP são capazes de compreender.