

Alta Disponibilidade: Confiabilidade e Alta Disponibilidade em Serviços de Rede

Resumo

Referência bibliográfica

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

O Planejamento de Redes de Alta Disponibilidade Confiáveis

As redes de computador têm como razão de sua existência a oferta de recursos e a conectividade entre os diversos usuários, e desta forma, para que os serviços sejam eficientes, entregues com qualidade e para que sejam feitas conexões estáveis, é necessário que tais redes sejam projetadas, desenhadas de forma que estes resultados sejam alcançados.

Sobre as redes, Soares et al. (2020) considera que são a soma de muitas partes e seu funcionamento correto promove a satisfação dos usuários. Neste contexto sua disponibilidade deve ser absoluta pois a demanda pelos serviços que oferece aumenta a cada dia. Desta forma a alta disponibilidade deve nascer com a rede, de forma que desde seu momento de ativação apresente desempenho otimizado, confiável, o que determina a consideração de diversos elementos, como o SPoF (Single Point of Failure - Ponto Único de Falha), e desta forma,

A disponibilidade é um dos pilares da segurança, que consiste em confidencialidade, integridade e disponibilidade. Dessa forma, falar em disponibilidade é praticamente o mesmo que falar em segurança, visto que uma não existe sem a outra. Sendo assim, a tolerância a falhas, que resulta em disponibilidade, traz uma maior segurança para as redes e os sistemas, garantindo que o sistema esteja disponível quando alguém precisar

utilizá-lo, resultando em uma maior confiabilidade. (SOARES ET AL. 2020, p.80).

Existe, porém, um ponto difícil de equalizar em um projeto de rede de alta disponibilidade: o atendimento deste nível de serviço em concorrência com a necessidade de se manter os custos em níveis aceitáveis e dentro do que prevê o projeto. No caso de provedores de serviços na nuvem os custos de implementação da alta disponibilidade são mais fáceis de serem incorporados ao projeto pois a empresa deve construir uma infraestrutura capaz de manter o SLA que oferece aos seus clientes, o que nos leva a algumas das características mais importantes das redes de alta disponibilidade:

Quadro 1: Características de redes altamente disponíveis

Eliminação do SPoF	O SPoF deve ser eliminado para garantir a redundância de operações.
Roteamento ininterrupto em torno de links com falha	O roteamento deve redirecionar sempre para o melhor caminho, para garantir disponibilidade e segurança.
Mecanismos rápidos de detecção de falhas (monitores)	Quanto mais rápida a detecção da falha e o redirecionamento, menor o risco de interrupção em sistemas e redes, resultando em uma maior confiabilidade.
Mecanismos rápidos de detecção de intrusões (monitores)	Uma intrusão pode causar muitos danos a uma rede e aos sistemas, como roubo de dados, inserção de artefatos maliciosos para manter o controle a distância.
Operações rápidas de substituição/correção	Quanto mais rápida a manutenção dos equipamentos, sistemas ou rede, mais rápido o serviço retorna sem maiores traumas, não perdendo em disponibilidade, visto que há a questão da redundância e dos demais elementos em conjunto.
Hardware e software confiáveis e sustentáveis	(alto MTTF [mean time to failure; ou tempo médio entre falhas] e baixo MTTR [mean time to repair, ou tempo médio para reparo]): esses elementos precisam ser confiáveis, a fim de necessitar o menos possível de trocas e manutenções.
Redundância versus custos incorridos pelos usuários:	Em geral, quanto maior a disponibilidade, maior é a redundância e, por consequência, maior o custo das soluções.

Componentes redundantes (NICs duplos, roteadores, etc.):	Por exemplo, clustering de roteador para WANs (Wide Area Network; ou Rede de Longa Distância, em português). Esses itens fazem a rede ficar disponível por mais tempo, trazendo, assim, uma maior confiabilidade e segurança.
Requisitos de largura de banda:	É preciso garantir que esses requisitos sejam atendidos por meio de modelagem ou outros cálculos sobre picos de tráfego, pico médio, variações sazonais, entre outros índices.
Operações, suporte, procedimentos de recuperação e treinamento para os operadores:	Os operadores e o suporte necessitam de treinamento sobre como devem agir em uma possível instabilidade, como necessidade de manutenção, substituição de hardware, parada total e demais incidentes passíveis de ocorrer.
Requisitos de SLA que cobrem os tempos de segurança, resposta e recuperação:	Esses itens normalmente são definidos em contratos de prestação de serviços e no (DRP), porém, para que sejam realistas, ao montar um plano desse tipo, este deve ser testado ao menos uma vez ao ano, a fim de garantir que a equipe esteja bem-treinada, bem como a funcionalidade real dos tempos e do plano de recuperação.
Documentação (design) e operações (runbooks):	Toda a rede e os sistemas necessitam de uma documentação completa, para que o conhecimento da estrutura e dos equipamentos facilite uma futura manutenção.

Fonte: Soares et al. (2020, p.80).

Além do investimento, o planejamento das redes de alta disponibilidade deve ser feito no nível lógico e considerar estruturas resilientes que forneçam bons recursos como o *load balancing* (LB), o balanceamento de carga. Com o LB uma rede mantém a distribuição de seus recursos de forma otimizada e com isso os diversos usuários acabam por ter serviços de melhor desempenho, mesmo em momentos de sobrecarga.

Alta disponibilidade nos Serviços de Rede

O funcionamento correto de uma rede altamente disponível não é algo facilmente percebido por seus usuários, pois consiste em uma série de recursos destinados a permitir que seus serviços

funcionem conforme foram projetados para. Assim, de forma geral e intrínseca os usuários de uma rede são dependentes da alta disponibilidade, pois seus sistemas são parte cada vez maior da sua rotina de trabalho. De acordo com Schmidt (2006, *apud* Soares et al. 2020, p.84),

[...] é difícil definir exatamente o que são os serviços de rede, porém podem significar serviços que devem estar disponíveis para uso em um aplicativo, mas que não fazem parte dele. Alguns exemplos de serviços de rede são: DHCP (Dynamic Host Configuration Protocol; ou Protocolo de Configuração Dinâmica de Host, em português), DNS, IP (Internet Protocol; ou Protocolo de Internet, em português), Active Directory, serviço de impressão, serviço de e-mail, entre outros.

Os principais serviços de uma rede, aqueles que são vitais ao seu funcionamento e por consequência atendem (mesmo que indiretamente), a todos seus usuários estão descritos no quadro a seguir:

Quadro 2: Principais recursos das Redes de Computadores

DHCP (Dynamic Host Configuration Protocol)	O DHCP é um protocolo de serviço TCP/IP (Transmission Control Protocol/ Internet Protocol; ou Protocolo de Controle de Transmissão/Protocolo de Internet, em português) utilizado em redes de computadores que atribui endereços IP automaticamente aos dispositivos no momento em que se conectam à rede. Desse modo, é possível a utilização de um endereço IP por apenas uma máquina de cada vez. Além disso, o DHCP distribui máscara de rede, gateway e DNS aos dispositivos.
DNS (Domain Name System)	DNS é um serviço onipresente em redes IP. Ele é responsável pelo mapeamento de nomes de host para endereços IP, e vice-versa. Sempre que é feita uma solicitação de acesso a um servidor (p. ex., www.google.com), o DNS mapeia esse nome para um endereço IP para que o acesso seja realizado. O DNS coloca nomes de host em domínios hierarquicamente estruturados e atribui a responsabilidade por domínios ou partes deles aos servidores DNS. Esses servidores formam um grande banco de dados distribuído de mapeamento de nomes.
AD (Active Directory)	O AD é um servidor de diretórios da Microsoft, o qual é implementado em protocolo LDAP, que permite o uso de um único diretório para o controle de acesso a todos os sistemas e serviços dentro de uma rede corporativa. Por possuir essa integração geral dos sistemas, um colaborador, por exemplo, não necessita criar um usuário e uma senha para cada sistema que tiver acesso, pois o AD cria credenciais únicas por usuário. (SOARES ET AL. 2020, p.93).

Fonte: Soares et al. (2020, p.85).

Portanto, a primeira visão sobre a importância dos serviços oferecidos em uma rede e a existência da alta disponibilidade está no que existe de mais fundamental nas redes, seus protocolos e funções, que primeiro vão permitir o correto funcionamento desta rede. Desta forma, para que uma rede seja altamente disponível, sua base lógica, com o DHCP, DNS, AD e demais protocolos, devem ser igualmente e altamente disponíveis.

O Processo de Monitoramento Alta Disponibilidade

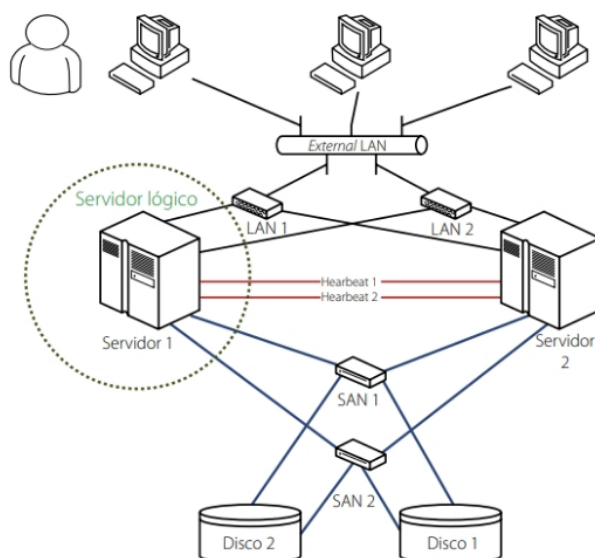
Em geral, uma rede de computadores requer bom gerenciamento para manter adequados níveis de performance. Quando a rede passa pelo planejamento de alta disponibilidade, o processo de gerenciamento, mais precisamente, seu monitoramento, passa a ser muito mais importante, crítico, pois passa a se tornar um elemento de alta disponibilidade.

Dentro dos elementos que devem ser construídos em um projeto de rede de alta disponibilidade existe a necessidade de se investir em monitoramento, seja via software ou hardware, para que exista um maior controle dos diversos fatores que podem comprometer a alta disponibilidade, de acordo com Soares et al. (2020, p. 96),

Para evitar as interrupções e buscar uma rede confiável e de alta disponibilidade, faz-se necessário investir em tecnologias e monitoramento de infraestrutura, visando minimizar ao máximo possíveis falhas. Há muitas formas para se minimizar ou evitar o SPoF, como a redundância de redes e servidores, o roteamento de links com falha, o balanceamento de carga e os mecanismos rápidos de detecção de falhas. A redundância é um dos itens mais importantes e mais utilizados, pois não existe alta disponibilidade sem redundância.

Quanto mais recursos uma rede precisa oferecer, maior será a complexidade de sua infraestrutura e assim, maior será a necessidade de robustos processos de monitoramento de seus diferentes aspectos. Temos diversos componentes que fazem parte de uma rede e que demandam cuidadoso monitoramento, como servidores, bancos de dados, a sua energia, seus diversos switches. A maioria destes componentes possui modelos com a opção de serem gerenciáveis, e desta forma oferecem dados aos sistemas de monitoramento. A figura a seguir apresenta um modelo de estrutura de rede de complexidade moderada:

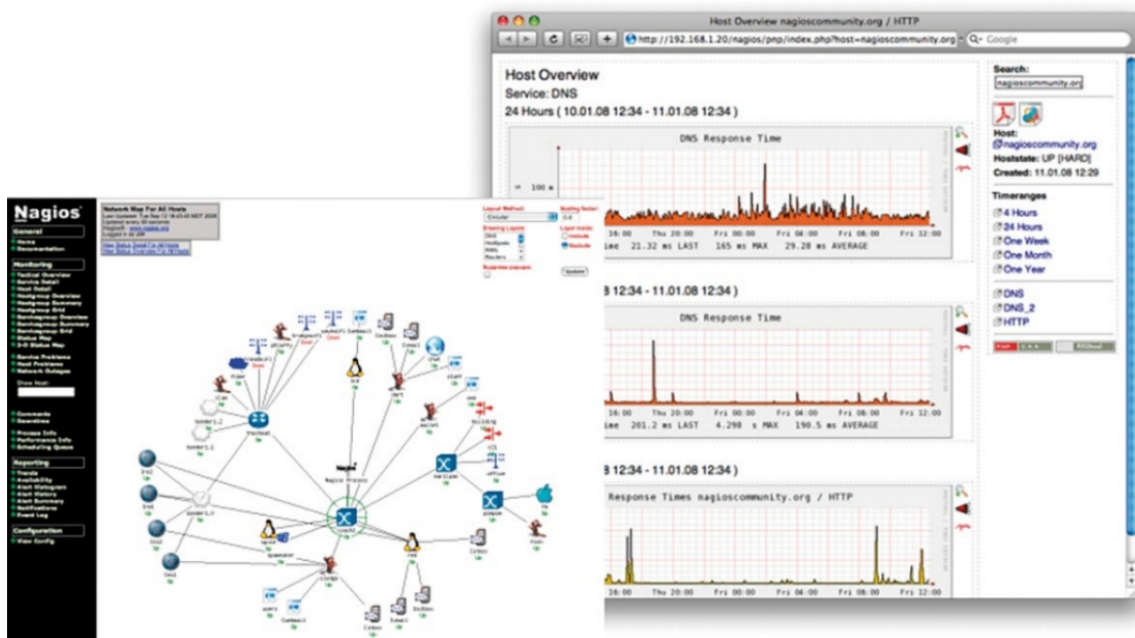
Figura 1. Cluster de alta disponibilidade



Fonte: Dantas (2005; p. 97).

O monitoramento representa à equipe que gerencia o TI a capacidade de se observar toda a rede e assim gera a capacidade de tomada de decisão na eventualidade da detecção de algum problema ou erro que eventualmente possa causar interrupções nos serviços. Este processo é realizado com o uso de software específico de monitoramento, que dentre outros indicadores, avalia o trânsito dos pacotes, conforme complementa Dantas (2005; p. 97), “Desse modo, é possível analisar de forma muito ampla incidentes no exato momento em que eles ocorrem, o que possibilita uma ação muito mais rápida e efetiva, além de permitir antever possíveis problemas que podem ocasionar a interrupção dos sistemas.” A figura a seguir apresenta um exemplo de tela de um sistema de monitoramento de rede:

Figura 2. Exemplo de tela de sistema de monitoramento.



Fonte: Furtunato (2020, documento on-line, p.98).

Internamente os sistemas de monitoramento usam o protocolo SNMP (*Simple Network Management Protocol* - Protocolo Simples de Gerência de Rede) que atua em redes TCP/IP e dever ser complementado por módulos de apoio a tomada de decisão que analisam indicadores e estatísticas da rede oferecendo soluções e indicações de ações assertivas.

Os Custos da Alta Disponibilidade de Redes

Devemos considerar os custos da alta disponibilidade nas redes em duas perspectivas: a da implementação e a que ocorre com a indisponibilidade. São diversos os processos elencados para que a composição do projeto permita uma consideração mais assertiva de seus custos, mas tudo começa com a definição do *budget*, da possibilidade de investimento que projeto pode receber.

Desta forma, em um primeiro momento podemos extrapolar os conceitos e dizer que a alta disponibilidade depende da disponibilidade de recursos financeiros para seu projeto de implementação, conforme apresenta Soares et al. (2020, p. 99),

Saber gerenciar os custos de um projeto é parte das atribuições do gestor de qualquer organização. No caso da alta disponibilidade, para definir quanto se deseja gastar para essa implantação, primeiro, deve-se conhecer quanto a organização perderá em momentos de interrupção da rede. Qualquer interrupção pode afetar a produtividade interna ou até mesmo causar a perda de credibilidade e de clientes que possam utilizar os sistemas da empresa.

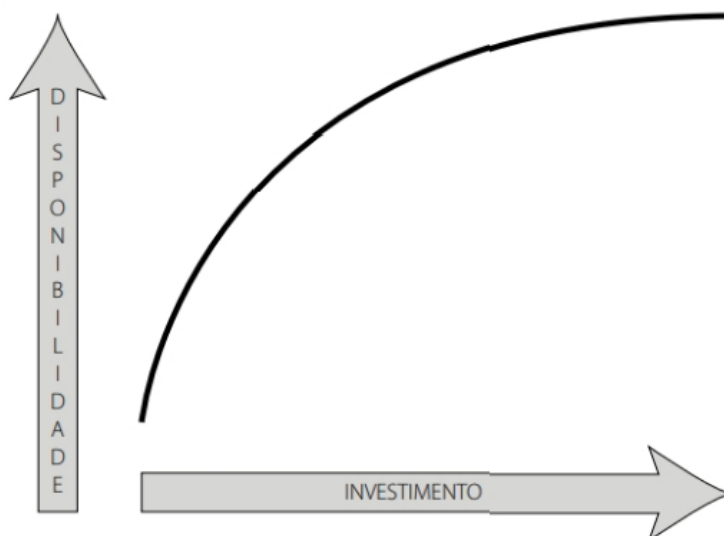
Existe grande dificuldade, na maioria dos casos, no processo de convencimento da necessidade de se investir em alta disponibilidade, mas com o passar dos anos, com a maior competitividade entre as empresas e uma maior procura por serviços de rede, online, um argumento vem sendo amplamente utilizado: o fato de que existe um custo muito menor da implementação da alta disponibilidade se comparado com o custo sofrido pelos momentos de inatividade da rede.

Portanto o investimento em alta disponibilidade deve ser visto como estratégico e como um elemento de longo prazo, para que seus custos se tornem vantagens a empresa e benefícios aos seus usuários, sempre mantendo em perspectiva, sua implantação, conforme Soares et al. (2020, p.99).

Implantar uma rede de alta disponibilidade custa caro, pois o ponto-chave da disponibilidade é a redundância. Imagine, por exemplo, quanto se gasta para projetar a rede. Com a redundância, o custo acaba sendo em dobro. Ou seja, quanto maior a disponibilidade, maior é o custo.

O custo da alta disponibilidade é considerável, e cresce conforme suas especificações, pois algumas redes e os sistemas que nelas estão inseridos, apresentam requisitos distintos de performance e tolerância de falhas, esta relação pode ser compreendida com a figura a seguir:

Figura 4. Disponibilidade versus investimento.



Fonte: Adaptada de Marcus e Stern (2003 *apud* SOARES et al., 2020, p.100).

Desta forma, é preciso provisionar recursos consideráveis para redes de maior complexidade, o que leva a compreensão de que a decisão por um projeto de rede de alta disponibilidade é estratégica e muitas vezes comercial, e desta forma devemos ressaltar que se considera não somente o montante a ser gasto mas o fato de que representa investimento de longo prazo.

Fazendo um paralelo, da mesma forma que veículos de alta performance, os superesportivos, dependem de uma série de sistemas e componentes igualmente feitos para alta performance, empresas com alta demanda dependem de redes altamente disponíveis. Em ambos os projetos existe um volume de recursos financeiros que deve ser provisionado, pois a alta disponibilidade, assim como a performance de um superesportivo somente são alcançados e mantidos com o uso de materiais, equipamentos e sistemas de performance superior.

Exercícios

1. Com a evolução tecnológica e computacional, as redes de computadores se tornaram ferramentas essenciais para a integração social e para a operacionalidade das organizações. E dada a importância da informação e o grande volume de dados que transitam pelas redes diariamente, desenvolver e implementar uma rede de computadores se tornou uma tarefa bastante complexa, e isso se deve algumas vezes pelas inúmeras demandas, que podem se tornar até mesmo conflitantes. Dessa forma, quais são as duas grandes demandas consideradas conflitantes no design de uma rede?

- a) Os custos e a alta disponibilidade.
- b) A alta disponibilidade e integridade.
- c) A Confiabilidade e integridade.
- d) A indisponibilidade e a Confidencialidade.
- e) A alta disponibilidade e a integridade.

2. A alta disponibilidade se tornou um elemento essencial para qualquer infraestrutura de TI, e representa um conjunto de componentes que estão interligados de forma a manter os serviços ininterruptos por determinado espaço de tempo. Para que seja possível manter esta alta disponibilidade de uma rede de, são necessárias determinadas características, como hardwares e softwares confiáveis, mecanismos de detecção de intrusão, eliminação dos SPoF entre outros.

Sobre as características de uma rede altamente disponível, qual a funcionalidade da eliminação dos SPoF (Single Point of Failure ou Ponto Único de Falha)?

- a) Redireciona para os melhores caminhos dos links com falhas.
- b) Garante a detecção de falhas.
- c) Promove uma maior rapidez na manutenção dos equipamentos.
- d) Garante a redundância das operações.
- e) Garante a confiabilidade do tempo médio entre falhas e tempo médio para reparo.

3. Os serviços de uma rede de computadores representam recursos que facilitam a operação de rede. De um modo geral estes serviços são de difícil definição, uma vez que estes recursos devem estar sempre disponíveis para sua utilização em um aplicativo, mas que efetivamente não fazem parte dele. Dentre os serviços de rede mais expressivos, estão os DHCP, Dynamic Host Configuration Protocol; ou Protocolo de Configuração Dinâmica de Host, DNS, IP (Internet Protocol; ou Protocolo de Internet, entre outros.

Dada a importância da disponibilidade destes serviços para que a rede se mantenha em operação, o que pode acontecer a uma rede de computadores na eventualidade da indisponibilidade do DHCP?

- a) Pode haver falhas na redundância.
- b) O usuário do sistema pode ter problemas na criação de credenciais.
- c) A ocorrência de conflitos de IP.
- d) Pode haver falhas na segurança da rede.
- e) Impossibilidade da realização de pesquisas inversas.

4. O desenvolvimento de redes altamente disponíveis, demanda uma série de recursos e planejamentos, pois para que esta rede se mantenha disponível exige que tenha como um dos requisitos principais, redundância. Dessa forma a criação de clusters, dispostos a preencher todos os requisitos de alta disponibilidade são ferramentas altamente eficientes.

Assim, compreendendo que um cluster compreende um sistema de computadores relacionados de forma a realizarem em conjunto uma tarefa, como deve ser a estrutura deste clusters de forma que ele seja considerado de alta disponibilidade?

- a) Não deve apresentar dispositivos em duplicidade.
- b) Deve apresentar apenas um servidor.
- c) Deve apresentar redundância em todos os pontos de sua estrutura.
- d) Dispensa o uso de ferramentas de monitoramento.
- e) Os computadores estão dispostos, sem a presença de um servidor.

5. Os sistemas de monitoramento de rede representam uma ferramenta essencial para que a equipe técnica tenha uma visão geral da rede como um todo, de forma que possíveis falhas possam ser antecipadas e que as soluções possam ser postas em prática o quanto antes, evitando assim possíveis episódios de indisponibilidade. Mas, para que este sistema de monitoramento seja ainda mais eficiente, é necessário que haja um sistema de tomada de decisão.

Neste contexto, qual a importância dos sistemas de tomada de decisão em um sistema de monitoramento para a manutenção da alta disponibilidade?

- a) O sistema de tomadas de decisão é capaz de determinar os custos de implementação de um sistema altamente disponível.
- b) Estes sistemas promovem a implementação de políticas de segurança e privacidade da rede.
- c) Estes sistemas auxiliam na determinação do sucesso ou a falha do ambiente através do monitoramento de falhas.
- d) O sistema de tomada de decisão disponibiliza dados para o reconhecimento da estrutura da rede.

- e) Este sistema promove a distribuição e carga de trabalho de acordo com as necessidades do sistema.

6. A alta disponibilidade de uma rede de computadores está diretamente relacionada com a redundância. A redundância, por suas características, possibilita que o sistema mantenha seu funcionamento, mesmo na ocorrência de falhas, o que é fator essencial para que não haja indisponibilidade. Entretanto, a implementação de um sistema que apresente redundância está diretamente ligada aos custos do projeto, uma vez que a redundância exige um aspecto diferenciado na estrutura da rede.

Em uma rede de computadores, como está relacionada a redundância com os custos de uma alta disponibilidade?

- a) Pelas simplicidades dos serviços de rede que oferecem redundância, os custos de implementação da alta disponibilidade são diminuídos.
- b) A estrutura da formação de clusters permite o uso de uma menor quantidade de equipamento, diminuindo os custos.
- c) A redundância eleva os custos da alta disponibilidade, pois exige a implantação de uma maior quantidade de equipamentos.
- d) A redundância eleva os diminui os custos de alta disponibilidade, pois dispensa a utilização de servidores.
- e) A necessidade de redundância diminui os custos de alta disponibilidade, pois não exige hardwares nem softwares sustentáveis.

Gabarito

1. Letra A.

Alternativa está correta pois existe uma demanda crescente de um sistema que ofereça alta disponibilidade, em contrapartida, sistemas altamente disponíveis possuem um custo de implementação consideravelmente mais elevado.

2. Letra D.

A alternativa está correta pois o SPoF (Single Point of Failure ou Ponto Único de Falha) representa um ponto que em caso de falhas, estas falhas podem prejudicar todo o sistema, afetando diretamente a redundância.

3. Letra C.

Alternativa está correta pois o DHCP atua no estabelecimento dos endereços de IP de uma rede, sem seu funcionamento novos computadores podem receber endereços que já existem o que causaria conflitos de funcionamento.

4. Letra C.

A alternativa está correta pois para que um cluster seja considerado de alta disponibilidade ele deve apresentar redundância em toda sua estrutura, ou seja, replicação de servidores, dados, switches, energia etc. Pois caso haja algum tipo de falha, o segundo servidor assume o serviço, evitando a indisponibilidade.

5. Letra C.

A alternativa está correta pois, um sistema de monitoramento de rede deve englobar um sistema de tomada de decisão que através de suas funcionalidades como, detecção e correlação de erros, roteamento de tráfego, entre outros, auxilia na determinação do sucesso ou da falha do ambiente.

6. Letra C.

Pois a redundância de um sistema exige uma quantidade maior de equipamentos, toda estrutura deve ter redundância, ou seja, estes equipamentos devem ser replicados, o que obviamente gera um maior custo de implementação.