

Sistemas Distribuídos



Anhanguera

AVALIE
SUA PROFISSÃO

QUANDO APARECER EM SEU
PORTAL UMA AVALIAÇÃO SOBRE
SEU CURSO, RESPONDA:



NOTAS

9 ou 10

SIGNIFICA QUE VOCÊ INDICA

NOTAS

7 ou 8

SIGNIFICA QUE VOCÊ NÃO INDICA



Anhanguera



Anhanguera



Segurança

Sem dúvida, um dos aspectos mais importantes no projeto de sistemas distribuídos é a segurança. Tipicamente, seja qual for a aplicação desenvolvida, sendo um sistema distribuído, funcionará em uma plataforma com várias máquinas, chamadas de nós, que replicam tal aplicação e, conforme já sabemos, a comunicação entre essas máquinas sempre ocorre por meio de redes de comunicação, tipicamente cabeadas.



A partir dessa análise, questões referentes à segurança desse sistema devem ser levadas em consideração. Segundo Coulouris et al. (2013), em termos de sistemas distribuídos, podemos pensar em dois níveis: o da confidencialidade e o da integridade dos dados.

- A confidencialidade dos dados refere-se ao acesso ao dado por indivíduos ou sistemas não autorizados.
- A integridade dos dados refere-se, além do acesso, à modificação do dado.



Obviamente a segurança é um tema altamente complexo, e existem várias disciplinas que abordam diferentes aspectos dessa interessante área de estudo, mas, em linhas gerais, o projeto de sistemas distribuídos em termos de segurança remete a um exercício de equilíbrio entre custo e ameaças (COULOURIS et al., 2013). Ainda conforme Coulouris et al (2013), os pontos de atenção em relação à segurança, no projeto de sistemas distribuídos, são:

- Portas são expostas: sistemas distribuídos são construídos com base em um conjunto de processos que oferecem serviços e compartilham informação. As portas de comunicação pelas quais esses serviços se comunicam são, intrinsicamente, abertas (para que clientes possam acessar tais serviços) e, dessa forma, um hacker pode enviar mensagem a qualquer uma dessas portas.



- Redes de computadores não são seguras: remetentes de mensagens podem ser falsificados, ou seja, um e-mail enviado por caique@caique.com pode não ter sido enviado pelo Caique; endereços IP podem estar duplicados, de forma que alguém malicioso possa receber as mesmas mensagens de um destinatário válido, etc.
- A validade das chaves criptográficas deve ser limitada: quanto mais tempo uma mesma chave estiver válida e ativa, maiores são as chances de esta estar comprometida, por ter maiores chances de ser conhecida (e explorada) por uma quantidade maior de pessoas e sistemas.



- Algoritmos de criptografia podem ter falhas: na atualidade, a melhor prática é de divulgar publicamente os algoritmos de criptografia para que a comunidade e entidades especialistas possam validar o algoritmo e sugerir melhorias, de forma que a privacidade esteja garantida pela chave criptográfica, e não pela inacessibilidade ao algoritmo utilizado.
- Hackers podem ter acesso a recursos poderosos: o custo dos recursos computacionais tem diminuído cada vez mais, de forma que máquinas poderosas estão acessíveis para a maioria da população. Assim, sempre considere que ataques podem ocorrer de inúmeras fontes e podem explorar vulnerabilidades utilizando inclusive ataques do tipo força-bruta (que tentam descobrir senhas por tentativa e erro, com simples “chutes”).



Escalabilidade

A escalabilidade é outro aspecto importantíssimo de um sistema distribuído. Escalabilidade é um termo comum em redes de computadores, e está intimamente ligada ao tamanho da rede. Segundo Tanenbaum e Steen (2008), um sistema cujo desempenho aumenta com o acréscimo de hardware e software, proporcionalmente à capacidade acrescida, é chamado escalável. É importante notar, entretanto, que um sistema dito escalável permite que se aumente ou diminua a quantidade de recursos. Você deve estar se perguntando: por que eu diminuiria a capacidade do meu sistema? Imagine a seguinte situação: você criou uma aplicação web que distribui conteúdo em vídeo para preparar estudantes para fazerem a prova do ENEM. Você roda essa aplicação, de maneira replicada, em um conjunto de servidores em nuvem de algum provedor de cloud computing conhecido do mercado, digamos, com 10 nós. Apesar da ideia ser excelente, você nota que a quantidade de usuários que utiliza sua plataforma cai drasticamente entre os meses de novembro e junho, uma vez que os estudantes começam normalmente a se preparar para esse exame – que ocorre anualmente, entre outubro e novembro – a partir de julho, quando estão de férias.



Anhanguera

Supondo que você paga para esse provedor de cloud computing R\$ 150,00 mensais para que este disponibilize os 10 nós de maneira contínua, não seria interessante que, nos meses de menor demanda, você diminuísse a quantidade de servidores para, por exemplo, metade e, assim, pague nesse período a quantia de, digamos R\$ 75,00 mensais? Nesse cenário, sua economia seria de R\$ 600,00, que você poderia investir em outros projetos. Esse é um exemplo típico de escalabilidade “para baixo”.

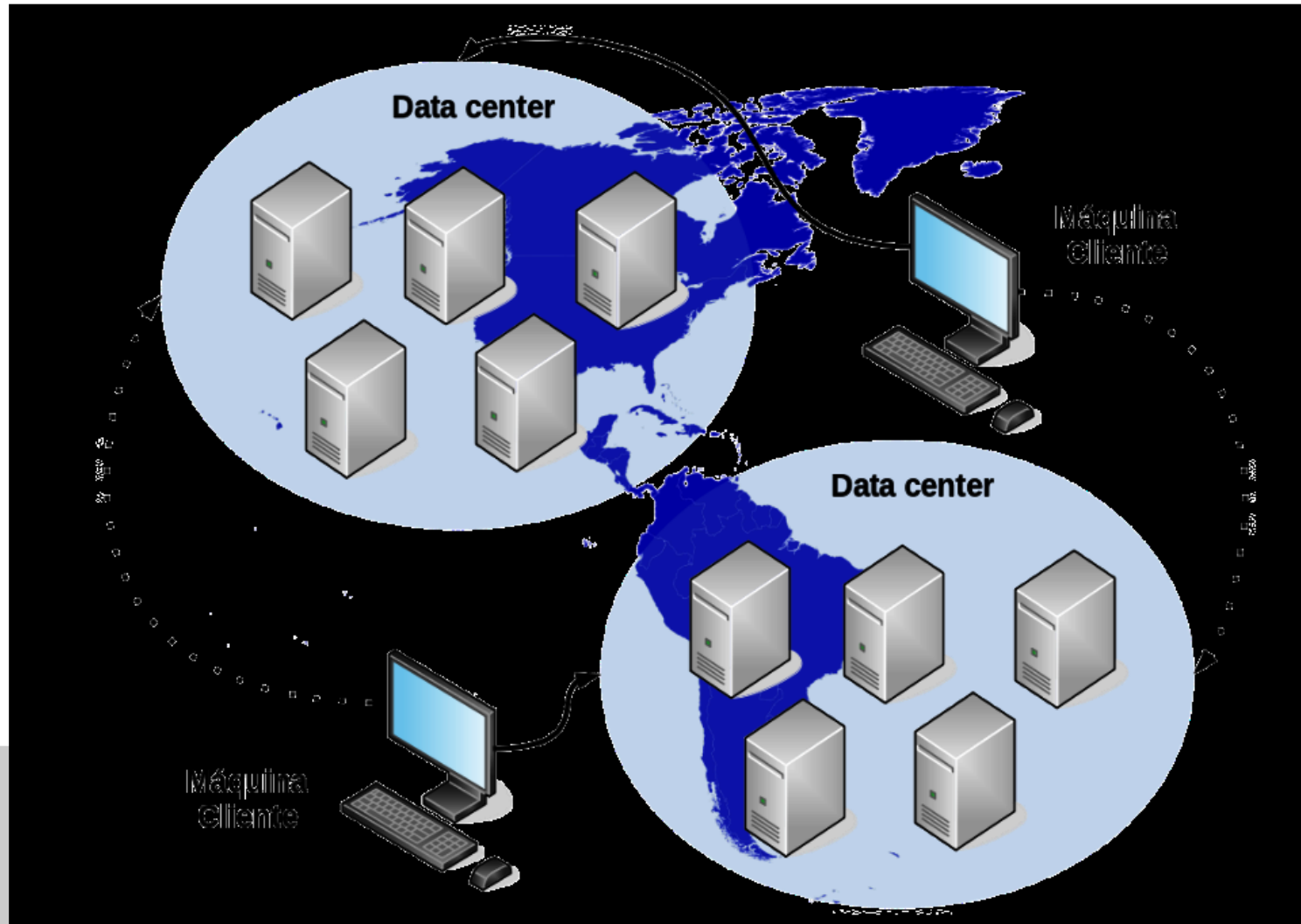


Anhanguera

Dois aspectos importantes a serem levados em consideração em relação à escalabilidade são em termos geográficos e administrativos. Escalabilidade em termos geográficos refere-se ao sistema que, apesar de apresentar-se como único para o usuário, está rodando em várias réplicas, em dois ou mais datacenters geograficamente distintos. Podemos, por exemplo, utilizar um determinado provedor de cloud computing que possua data centers no Estado de São Paulo, aqui no Brasil, e no Estado do Arizona, nos EUA. A Figura 2.8 ilustra tal cenário.



Anhanguera





Anhanguera

O benefício desse tipo de configuração é fornecer uma melhor experiência – em termos de conectividade e latência (atrasos na rede) para os usuários, uma vez que os usuários mais próximos do hemisfério Norte podem acessar a aplicação através dos data centers nos EUA, e os usuários do hemisfério Sul podem acessar a aplicação através dos data centers no Brasil. Outra vantagem é que, na ocorrência de um desastre, por exemplo, um furacão passar por Arizona, que comprometa o data center, todos os usuários poderão acessar o data center de outra localidade, incluindo os usuários do mais próximos ao data center atingido (ainda que a usabilidade, do ponto de vista desses usuários, seja ligeiramente comprometida, devido à maior distância desse data center).



Anhanguera

Em termos administrativos, escalabilidade refere-se ao escopo administrativo, que é afetado pela escalabilidade geográfica e é um conceito bastante simples de ser compreendido, embora muitas vezes ignorado. Imagine que, no cenário da Figura 2.8, os links de comunicação do lado dos EUA sejam fornecidos por provedores de Internet daquela região, ao passo que os links de comunicação no lado do Brasil sejam fornecidos por provedores de Internet daqui. Caso o link no lado dos EUA fique indisponível, não vai adiantar entrar em contato com o provedor de Internet daqui do Brasil, pois é uma empresa diferente da que fornece o serviço nos EUA, administrativamente falando. Ou seja, o escopo administrativo foi, inerentemente, ampliado, o que significa que o responsável pelo sistema distribuído terá mais trabalho para administrá-lo, incluindo, por exemplo, a necessidade de abrir um chamado de suporte técnico em outro idioma.



Anhanguera

Tolerância a falhas / Resiliência

Quando falamos de aspectos no projeto de sistemas distribuídos, temos que falar de resiliência de processos, esse aspecto é também um dos principais objetivos a serem atingidos em uma aplicação distribuída. A resiliência de processos está relacionada com o sistema ter uma comunicação confiável entre as camadas de Cliente e Servidor. Sua ideia básica é de que os seus processos da nossa aplicação sejam replicados em grupos, isso faz com que o sistema tenha uma proteção contra falhas relacionadas a processos.



Anhanguera

Para conseguir criar projetos tolerantes a falhas temos que ter em nossos sistemas uma detecção de falhas, assim como conseguir mascarar todas as falhas apresentadas e a replicação de nosso sistema, para que ela seja imperceptível. Só podemos atingir estas características de acordo com questões de projetos, e a mais importante neste caso é verificar qual o grupo de processos que a nossa aplicação deverá conter. Para isso, vamos entender o que são os grupos Simples e grupos Hierárquicos, e como esses processos são associados a esses grupos.



Anhanguera

Quando falamos de resiliência de processos, organizamos em grupos os processos que são considerados idênticos. Portanto, quando uma mensagem é enviada a um grupo de nosso sistema, a ideia é que todos os processos membros desse grupo recebam a mensagem. Se ocorrer uma falha em um dos processos no tratamento da mensagem, outro processo desse grupo deve tratar a mensagem no lugar do processo com falha.

Em um sistema distribuído temos grupos dinâmicos, ou seja, os processos podem se mover entre os grupos que são criados. Um processo de um cadastro de usuário do site XYZ pode enviar uma mensagem para a realização do cadastro a um grupo de servidores sem precisar quantos existem e quem eles são.



Anhanguera

Grupo Simples X Grupo Hierárquico

Quando falamos de grupos Simples, todos os processos são iguais e todas decisões são tomadas entre todos os processos, ou seja, de forma coletiva. A grande vantagem deste tipo de grupo é que não há um ponto único para falha. Mesmo que ocorra falha ou caia algum processo, o grupo continua mantendo o serviço em funcionamento. Portanto, a execução do sistema não é centralizada a um só ponto.



Anhanguera

A desvantagem que encontramos nesse tipo de grupo é que a tomada de decisão tende a demorar mais pois cada decisão deve ser priorizada pelos processos, tendo assim uma votação antes da tomada de decisão.

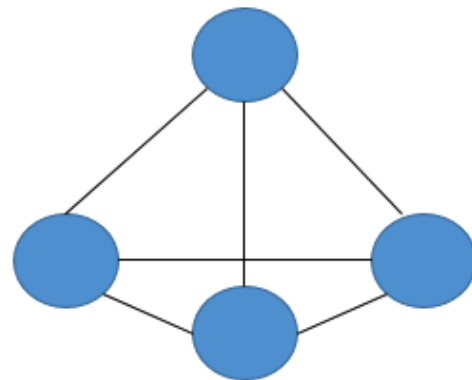
Já os grupos Hierárquicos, como o nome mesmo já diz, são baseados em uma hierarquia, portanto existem processos considerados mais importantes e que controlam toda execução. Nesses grupos temos um processo chamado “coordenador” e os demais chamamos de “operários”. Sempre que chega uma nova requisição no sistema ela é enviada ao processo coordenador, que decide o melhor operário para executá-la.



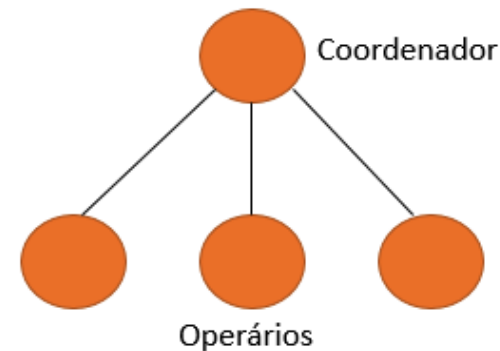
A grande vantagem desse tipo de grupo é que as decisões são centralizadas, portanto temos mais agilidade na tomada de decisão, o que gera um retorno mais rápido. Já a grande desvantagem apontada é que, caso ocorra uma falha no processo coordenador, o serviço todo para.

Podemos observar na Figura 2.9 a estrutura dos grupos apresentados e como seus processos se comunicam.

Grupo Simples



Grupo Hierárquico





Anhanguera

Podemos observar que, na comunicação simples, os processos se comunicam entre si e decidem em conjunto qual o processo mais adequado para executar determinada ação durante o funcionamento de um sistema. Já na comunicação hierárquica os processos estão divididos entre “coordenador” e “operários”, em que o processo coordenador irá definir o processo operário mais adequado para executar determinada ação durante o funcionamento de um sistema.



Anhanguera

Heterogeneidade

O termo heterogeneidade vem de heterogêneo, ou seja, algo desigual, que apresenta estrutura, função ou distribuição diferentes. Portanto, aplicando isso ao conceito de computação, quando temos heterogeneidade estamos falando de um sistema que contenha em sua composição máquinas (nós) de sistemas operacionais, recursos (hardware) e até mesmo fabricantes diferentes. Esse é um dos aspectos mais frequentes de um sistema que utiliza arquitetura distribuída. Geralmente, o sistema é composto por máquinas de diversas características diferentes que se comunicam para manter o funcionamento.



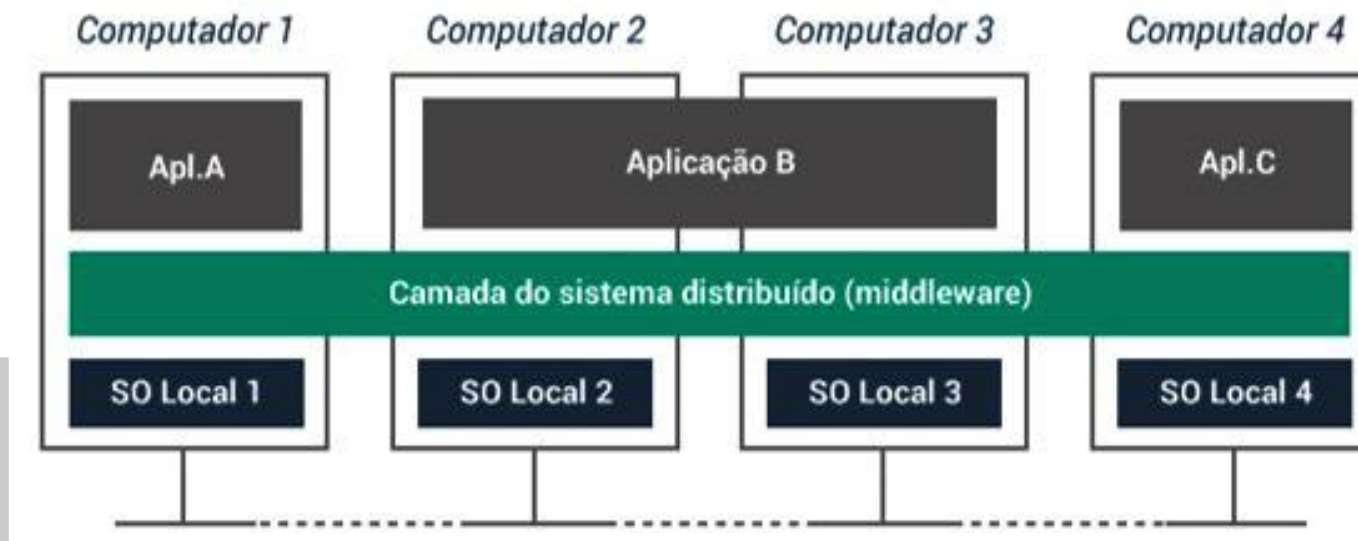
Anhanguera

Os protocolos de redes são fundamentais para que essa comunicação entre máquinas diferentes ocorra, porém, na maioria das aplicações distribuídas, precisamos também de um middleware para que essa comunicação ocorra.

O middleware pode ser considerado um conjunto de padrões e funcionalidades que atua como uma camada central entre a nossa plataforma, o sistema operacional e as nossas aplicações. Essa camada central permite que em um sistema distribuído rodem diferentes aplicações em diferentes plataformas e que todas elas consigam se comunicar adequadamente.



Na Figura 2.10 temos um exemplo da atuação da camada Middleware perante diferentes computadores (máquinas), aplicações e sistemas operacionais. Podemos observar que a camada middleware liga aplicações A, B e C com os sistemas operacionais 1, 2, 3 e 4. Essas aplicações e sistemas operacionais estão em diferentes computadores, 1, 2, 3 e 4. Estes irão interagir entre si para a execução de um sistema distribuído, e isso só é possível devido à camada do sistema distribuído, chamada middleware.





Existem vários frameworks utilizados para implementação de plataformas de middleware, e os mais conhecidos são:

- CORBA - Framework de implementação de middleware baseado na linguagem de programação C++.
- .NET Remoting - Framework de implementação de middleware, baseado no ambiente de programação .NET, que pode utilizar várias linguagens diferentes, por exemplo: Visual Basic, C#, entre outras.
- Akka.NET - Framework de implementação de middleware, baseado no ambiente de programação .NET, que pode utilizar várias linguagens diferentes, como por exemplo: Visual Basic, C#, entre outras.
- Java RMI - Framework de implementação de middleware, baseado na linguagem de programação JAVA.
- JAX-WS - Framework de implementação de middleware, baseado na linguagem de programação JAVA.

Daqui a algumas seções vamos utilizar o JAVA-RMI para criar alguns exemplos.



Anhanguera

Quando planejamos implantar um sistema distribuído, há vários fatores importantes que devem ser levados em consideração. Podemos chamá-los de aspectos de projeto, e, no projeto de sistemas distribuídos, os principais são: Segurança, Escalabilidade, Resiliência e Heterogeneidade.

A capacidade de máquinas com diferentes sistemas operacionais se comunicarem na execução de um sistema se relaciona a qual aspecto de projeto?

- a) Segurança.
- b) Escalabilidade.
- c) Resiliência.
- d) Heterogeneidade.
- e) Transparência.



Um dos aspectos mais importantes no projeto de sistemas distribuídos é a segurança. Tipicamente, seja qual for a aplicação desenvolvida, sendo um sistema distribuído, esta funcionará em uma plataforma com várias máquinas, chamadas de nós, que replicam tal aplicação. Essa plataforma tem a comunicação entre as máquinas feita através de redes de comunicação.

Sabendo que a segurança é um aspecto muito importante, analise as afirmações abaixo e escolha a opção com pontos de atenção em relação à segurança do sistema.

I – Máquinas de diferentes sistemas operacionais se comunicam.

II – Portas abertas.

III – Algoritmos de criptografia podem ter falhas.

IV – Camada de Middleware.

V – A validade das chaves criptográficas.

- a) Somente as afirmações I e IV são pontos de atenção em relação à segurança.
- b) Somente as afirmações I e II são pontos de atenção em relação à segurança.
- c) Somente as afirmações II, III e V são pontos de atenção em relação à segurança.
- d) Somente as afirmações II e V são pontos de atenção em relação à segurança.
- e) Somente as afirmações I e III são pontos de atenção em relação à segurança.