

Redes e Sistemas Distribuídos

Aspectos de projeto dos sistemas distribuídos

Você sabia que seu material didático é interativo e multimídia? Isso significa que você pode interagir com o conteúdo de diversas formas, a qualquer hora e lugar. Na versão impressa, porém, alguns conteúdos interativos ficam desabilitados. Por essa razão, fique atento: sempre que possível, opte pela versão digital. Bons estudos!

São vários os aspectos de projeto em sistemas distribuídos, alguns dos principais são: a segurança, escalabilidade, resiliência e heterogeneidade. Nesta webaula vamos explorar com maior detalhamento sobre segurança e escalabilidade.

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, esta funcionará em uma plataforma com várias máquinas, chamadas de nós, que replicam a 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 a 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.

Confidencialidade

A confidencialidade dos dados refere-se ao acesso ao dado, por indivíduos ou sistemas não autorizados.

Integridade

A integridade dos dados refere-se à quando, além de ser acessado, o dado foi modificado.

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).

Pontos de atenção

Ainda conforme Coulouris *et al.* (2013), os pontos de atenção em relação a 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 nas 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

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, certifique-se de considerar que ataques podem ocorrer de inúmeras fontes, e que podem explorar vulnerabilidades utilizando inclusive ataques do tipo força-bruta (que tentam descobrir senhas por tentativa e erro, através de simples “chutes”).

Escalabilidade

A escalabilidade é outro aspecto importantíssimo de um sistema distribuído. Escalabilidade é um termo comum em termos de 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 o mesmo aumente ou diminua a quantidade de recursos.

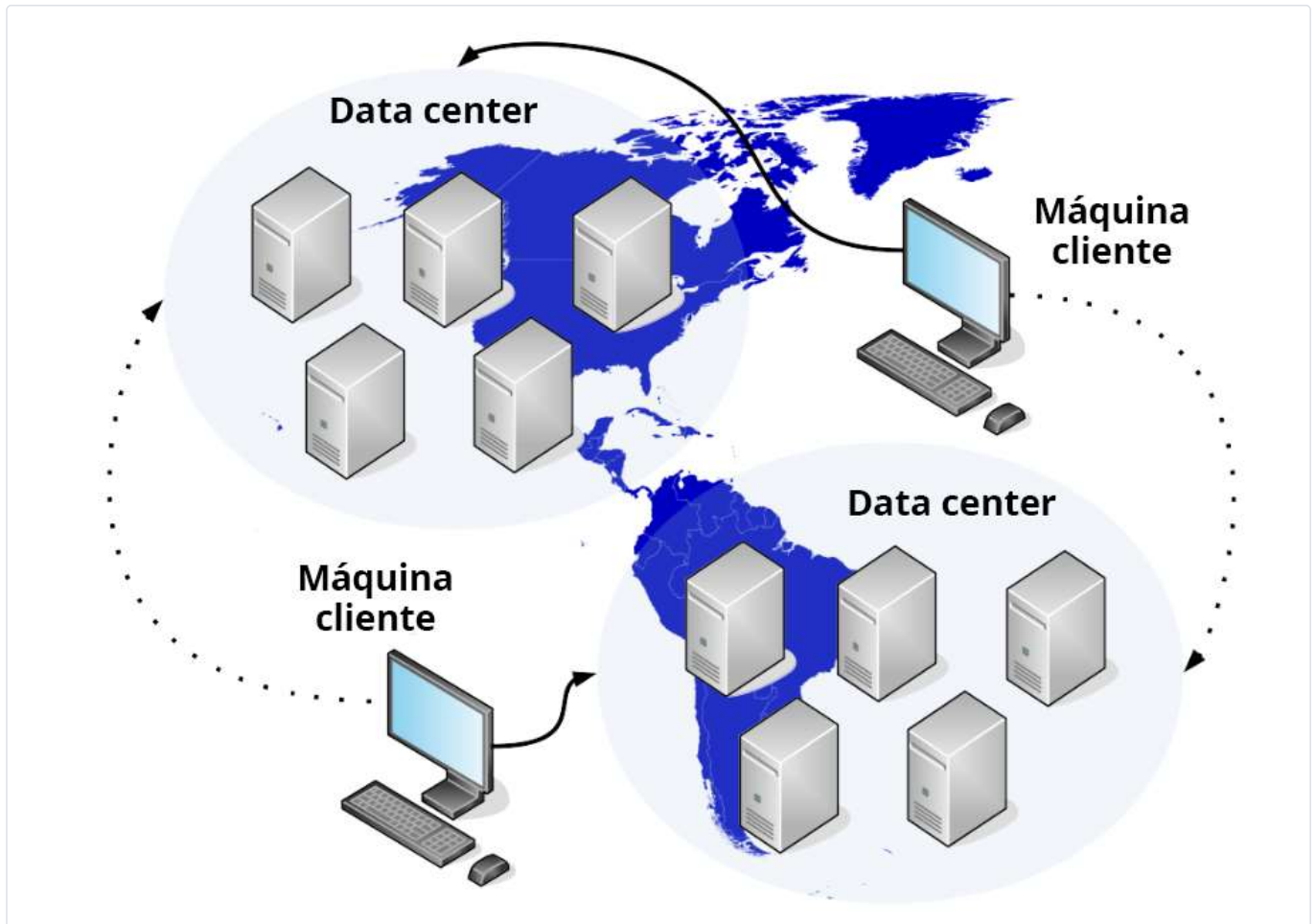
Diminuição da capacidade do sistema

Imagine a seguinte situação: você criou uma aplicação web que distribui conteúdo em vídeo para preparar estudantes a fazer 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. 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”.

Dois aspectos importantes de 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.

Exemplo de escalabilidade geográfica



Fonte: elaborada pelo autor.

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, de um furacão passar por Arizona, e comprometer aquele data centers, todos os usuários poderão acessar o data centers de São Paulo, incluindo os do hemisfério Norte (ainda que a usabilidade, do ponto de vista dos usuários do hemisfério Norte, seja ligeiramente comprometida, devido a maior distância desse data centers).

Escalabilidade em termos administrativos 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 imagem de escalabilidade geográfica, que estudamos anteriormente, 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.

Estudamos nesta webaula sobre escalabilidade e segurança. Para aprender sobre os demais aspectos, consulte o livro didático.

Para visualizar o vídeo, acesse seu material digital.