

## Visão geral

### O que é um sistema distribuído?

Um sistema distribuído é aquele que é definido como um conjunto de unidades de processamento independentes, que através da troca de comunicação e gerenciamento de sincronização pode processar uma aplicação em diferentes localidades em sistemas com características próprias diferentes, dando a impressão ao usuário que toda a aplicação é gerenciada por um sistema único. Quando falamos em sincronização, temos o conceito de sincronização em um sistema centralizado e no sistema distribuído. No sistema centralizado a sincronização é feita através do compartilhamento de áreas de memória, já no sistema distribuído ocorre a sincronização através da troca de mensagens. A aplicação no sistema distribuído pode ser dividida em “partes” diferentes e ser processada em diversos núcleos de processamento.

### Qual o objetivo de um sistema distribuído?

O objetivo é criar a ilusão que a aplicação (ou as aplicações) estão sendo processadas em um único sistema, permitindo a sensação que tudo isso ocorre sem o compartilhamento de áreas de memória, no entanto, a sincronização é feita a partir de trocas de mensagens. Faz parte do objetivo a situação da aplicação ser processada de modo que o ambiente que opera forneça situações favoráveis ao compartilhamento de recursos, sabendo que diferentes recursos estarão disponíveis em unidades de processamento diferentes.

### Exemplo de um sistema distribuído:

A própria internet, afinal de contas a internet é um exemplo claro disso tudo, pois temos um conjunto de computadores independentes que interligam-se mutuamente com o objetivo de compartilhar recursos, dando a impressão ao usuário que isso tudo é um único sistema.

### Características de um sistema distribuído:

Primeiramente é importante ter a visão de que uma aplicação pode ter diversos tipos de complexidade possível, não apenas com relação ao custo computacional mas também sob aspectos de gerenciamento e necessidade de recursos, ou seja, uma aplicação pode ser sincronizada entre alguns poucos computadores em uma pequena rede (até mesmo uma rede local) ou até mesmo aplicações que necessitem de um alto grau de sincronizações, como a utilização global da própria internet.

Otimização do compartilhamento de recursos: Isso significa estar apto para compartilhar com desempenho e segurança recursos físicos ou lógicos, como por exemplo, impressoras, scanners, dados, espaço em disco, processamento entre outros. Tudo isso é gerenciado por um software servidor.

É dever do sistema distribuído refletir a estrutura organizacional que é pretendida a ele, isso significa que deve estar muito claro e declarado o objetivo da aplicação e os recursos com núcleos de processamento que serão dispostos.

Talvez uma das características mais marcantes e presentes é a transparência, eu arrisco dizer que ela anda de “mãos dadas” com o compartilhamento otimizado de recursos, pelo fato de que se o usuário perceber de alguma forma que o sistema é gerenciado e processado por diversas unidades de processamento, já falhou o objetivo do sistema distribuído, que é dar a ilusão ao usuário que toda a aplicação é processada em um sistema único. Quando falamos de

transparência, é importantíssimo lembrar que da mesma forma que é um conceito importante e vital, ela foi dividida em diversas categorias.

#### Transparência de acesso:

Quando nós temos a possibilidade de acessar e/ou interagir com um dado, recurso ou até mesmo um dispositivo, nós precisamos de um meio para acessar. A transparência nesse aspecto significa que a forma pela qual eu interajo com a possibilidade de acessar e também a transformação desse acesso em utilização, faça parte de um único sistema.

#### Transparência de localização:

Os recursos, dados e dispositivos que serão compartilhados e acessados, não precisam e não devem conter nenhuma informação relevante que forneça ao usuário a localização região que esse recurso está contido. Por exemplo, se eu quiser acessar um arquivo, um documento que fale sobre a utilização da água, para o usuário, não deve ser relevante o fato de esse documento estar localizado no Brasil ou na Austrália, pois o importante é que o usuário tenha acesso ao documento, sem se importar com a localização física do recurso.

#### Transparência de migração:

Eu sei que na internet, nós temos muitos recursos compartilhados, é bastante comum que pela própria turbulência de acessos, seja necessário que esse recurso seja alocado para outro servidor, por exemplo, mas isso não deve interferir com a maneira que eu tenho acesso ao arquivo. Os softwares gerenciadores devem ser aptos o suficiente para referenciar corretamente os arquivos e recursos em caso de migração.

#### Transparência de replicação:

É bastante comum na internet, os recursos e dados serem tratados como objetos, isso significa que a instanciação deles é também muito usual na própria rede. Por isso é bastante interessante e necessário que o sistema distribuído seja capaz de gerenciar a replicação de informações de maneira transparente ao usuário.

#### Transparência em relação à falhas:

Esse é um fator de extrema importância, uma vez que temos a possibilidade de compartilhar recursos, não é nada agradável e necessário que os problemas também sejam compartilhados, por isso, em caso de falhas é essencial que o sistema seja capaz o suficiente para gerenciar sem que isso passe a espalhar para todo o sistema de compartilhamento, ou seja, os outros usuários podem continuar utilizando a aplicação sem compartilhar com essa falha também, sendo assim um problema considerado isolado.

Um sistema distribuído precisa necessariamente ser confiável, isso significa ter condições de acessar uma informação ou recurso que não esteja prejudicialmente modificado, com as suas devidas permissões de acesso.

O sistema também precisa ser tolerável contra falhas, isso significa a necessidade de possuir backups que possam ser ativados em tempo de operação, incluindo redundâncias de hardware até mesmo do software, no qual replicar informações é fundamental para que você possa garantir que um dado que teve falha passe a ser substituído em tempo suficientemente tolerável para que seja aproveitado da forma correta.

O desempenho também é fator importante nos sistemas distribuídos, por exemplo, podemos ter uma aplicação que compartilha diversos núcleos de processamento, é necessário que a carga do processamento esteja devidamente distribuída entre os componentes que estão efetuando o processamento, isso é vital para o sistema distribuído.

Além disso, o sistema deve possibilitar a escalabilidade, através da inclusão de novos recursos, dispositivos, sem que o desempenho e integridade atual sejam afetados. Isso significa que você deve estar preparado para possíveis necessidades de melhorar o sistema, sem que o processamento atual seja denegrido. Isso significa agregar recursos, seja de software e hardware possibilitando também o isolamento de falhas.

Outro fator preponderante é a flexibilidade, pois de nada adianta possuir um sistema distribuído que seja „imutável“, portanto é necessário que seja possível adaptar novos recursos e alterar o projeto atual, sem perder integridade, desempenho e segurança para o processamento atual.

Também é importante e é uma característica do sistema distribuído a não-necessidade do controle global dos processos, pois os recursos são utilizados de forma concorrente e a sincronização é feita por troca de mensagens, sem a necessidade de um controle global do sistema, de forma centralizada.

Sabendo que sempre existem vantagens e desvantagens para serem consideradas, citarei primeiro as vantagens, que são compartilhamento de recursos, ganho de desempenho, comunicação e confiabilidade.

O compartilhamento de recursos pode ser desde recursos físicos, como impressoras, até mesmo recursos lógicos, como arquivos e até mesmo acesso à um banco de dados concentrado.

A velocidade da computação pode ser incrementada de forma que existe o processamento concorrente de tarefas, que ao invés de ser processado em um núcleo único, ele pode se dividido em “n” processadores para reduzir o tempo de operação e ganho de disponibilidade.

Agora falando um pouco das desvantagens, citarei o caso de um sistema distribuído que envolva o critério de escalabilidade, por exemplo, com as melhorias e avanço tecnológico com tendências de migrar recursos e serviços para a internet, podem ocorrer situações externas que gerem um problema ou desconforto, como por exemplo o fato da prestadora de serviços de rede / internet estar sobrecarregada ou indisponível temporariamente. Sabendo também das tendências de tecnologia é sempre esperado que ocorram mudanças benéficas na estrutura e planejamento estratégico de tecnologia, talvez seja necessário modificar a interface de hardware e até mesmo de rede que agreguem custos bastante significativos, tratando-se de uma empresa.

Existem também problemas com relação aos softwares gerenciadores do sistema distribuído, uma vez que é uma tendência relativamente nova, os recursos tecnológicos de hardware, plataforma de software, sistema operacional e integração com a aplicação, podem resultar oscilações sabendo que é uma plataforma que não é 100% consolidada e também está em constantes melhorias e avanços.

Dificuldades para negar o acesso indevido é um fator crítico, sabendo que uma das características do sistema é ter múltiplas instâncias de processamento, isso significa que uma aplicação pode ser gerenciada por diversas localidades, dando uma vazão mais pertinente à acessos concorrentes, isso significa como se fosse uma grande balança, ao mesmo tempo que podemos ter um grande grau de disponibilidade, podemos também ter um grau muito grande de possibilidades de acesso, no qual o risco de acesso não autorizado cresce proporcionalmente.

Os sistemas distribuídos podem ter planejamento específico e a interface que será utilizada, darei agora o exemplo de uma rede de computadores. No qual o meio de acesso é a rede local, através de protocolos específicos. Sempre se associa um recurso à um objetivo pré estipulado.

Situações que podem prejudicar o desempenho do sistema, podem ser definidas como, por exemplo: Ter um servidor com altíssimo grau de processamento e velocidade de transmissão, com estações de trabalho de baixo processamento, a sensação que se dá é que não está sendo usado todo o processamento que poderia ser alocado ao uso do sistema, o oposto também é válido, se tiver as estações com muita capacidade de processamento e servidor com baixo processamento e velocidade de transmissão, o tempo de resposta é afetado negativamente. Caso tenhamos excelentes servidores, com excelentes estações de trabalho, mas um meio de transmissão de baixa capacidade, o tempo de resposta ficará degradado e demorado, portanto a sensação de falta de desempenho é muito alta.