

Redes e Sistemas Distribuídos

Redes de Computadores e Seus Protocolos

Introdução a Redes De Computadores

Desenvolvimento Histórico

Ao termino da Segunda Guerra Mundial e o começo da Guerra Fria, os embates tecnológicos e econômicos entre Estados Unidos e União Soviética se tornaram intensos. Nesse período o então presidente americano Eisenhower incentivou pesquisas voltadas a redes, que ficou conhecida como ARPANET. Inicialmente o objetivo era compartilhar dados entre bases militares e universidades para, caso a União Soviética atacasse uma dessas bases, os dados não seriam perdidos, pois teria uma espécie de BACKUP em outros pontos. Essa tecnologia conectava apenas quatro pontos dentro dos EUA. Dentro de 5 anos a ARPANET dentro dos EUA já conectavam mais de 200 pontos e começou também a contar com pontos na Inglaterra. Em meados de 1986, a ARPANET atingiu a marca de 100.000 computadores conectados e espalhados pelo mundo todo. Nessa época surgiam também pequenas redes comerciais incluindo pessoas que não eram mais universitários nem militares. Com isso, o Governo Dos Estados Unidos que inicialmente controlava a rede, percebeu que não seria mais possível fazer esse controle principalmente devido ao seu tamanho que chegou a nível global e decidiu então juntar todas essas pequenas redes e formar assim a INTERNET.

Arquiteturas

Na Arquitetura **Cliente e Servidor** existe um modelo de comunicação distribuída baseado em redes de computadores com servidores provendo acessos e controle aos sistemas e clientes, chamados de estações de trabalho, que se conectam aos servidores para acessos aos recursos de rede e dados. O papel bem definido do servidor é manter a aplicação e seus dados à disposição dos clientes. Na Arquitetura **Ponto a Ponto** as redes não tem um servidor nem um administrador bem definido. Todo mundo pode ser servidor e cliente ao mesmo tempo e não há uma maquina com poder computacional suficiente para gerenciar os demais dispositivos da rede. Redes domésticas podem ser exemplos de redes com a arquitetura Ponto a Ponto. Esse tipo de rede tem um custo mais baixo devido a simplicidade e a baixa segurança.

Abrangência Geográfica - Escala

PAN (Personal Area Network) - Conhecidas como Redes Pessoais, são redes de pequena abrangência geográfica que permitem que os dispositivos se comuniquem sem a necessidade de um intermediário como um roteador. Um exemplo seriam as conexões feitas por Bluetooth. Geralmente encontrados em residências.

LAN (Local Area Network) - Conhecidas como Redes Locais é uma rede particular que opera dentro de um espaço físico limitado, como uma residência, um escritório ou uma empresa. São muito usadas para conectar computadores pessoais e aparelhos eletrônicos permitindo que compartilhem recursos e troquem informações.

CAN (Campus Area Network) - Quando diversas Redes Locais se conectam, formando uma única rede. Geralmente ocorre em universidades que possuem muitos prédios e edificações com Redes Locais para cada uma delas, mas que quando se conectam forma uma CAN, que é uma rede para todo o campus da universidade.

MAN (Metropolitan Area Network) - São redes de comunicação que abrangem áreas de maior extensão. Geralmente usada para conectar fisicamente a rede de uma cidade a outra ou a sede de uma empresa às suas filiais em lugares distantes.

WAN (Wide Area Network) - Essa é uma classificação de rede a nível global. O Maior exemplo de WAN que podemos citar é a internet que conecta o mundo inteiro.

Acesso

Intranet - É uma rede privada e interna em uma organização ou empresa, com acessos restritos a usuários e dispositivos homologados.

Extranet - É uma rede que abrange sites corporativos com informações internas e acessos geograficamente externos. A Internet como rede global de computadores é uma estrutura de extranet com acesso abrangente.

Topologias

Topologia em Malha - Em uma topologia de malha, a maioria dos dispositivos se conecta diretamente. Isso oferece vários caminhos para entrega de dados. Os dados são entregues pela distância mais curta disponível para transmissão. Nessa topologia, a principal vantagem é ter uma conexão direta entre cada dispositivo. Tendo como principal desvantagem a complexidade das conexões.

Topologia em Barramento - Orienta os dispositivos ao longo de um único cabo que vai de uma extremidade da rede à outra. Os dados fluirão ao longo do cabo conforme ele se desloca até seu destino. Nessa topologia para redes locais, a vantagem econômica é o uso de um único cabo para ligação, porém essa ligação também é vista como desvantagem uma vez que a interrupção desse cabo único representa a paralização total da rede.

Topologia em Anel - Modelo em que cada dispositivo tem uma conexão direta e dedicada com outros dois hosts. Quando um host recebe um sinal destinado a outro, seu repetidor regenera os dados e os encaminha para o destino. A principal vantagem é a facilidade de instalação, a desvantagem é que os dados são transmitidos em sentido único.

Topologia em Árvore - Topologia em que os hosts estão organizados abaixo de dispositivos de rede formando ramificações de elementos. A vantagem desse tipo de topologia é a organização da estrutura, o controle de hosts e o gerenciamento da rede e como desvantagem, existe a necessidade de se prover sistemas redundantes para que a rede não seja prejudicada quanto a falhas em dispositivos.

Topologia em Estrela - A rede é organizada de forma que os hosts sejam conectados a um HUB central, que atua como um servidor. Qualquer dado enviado pela rede viaja primeiro pelo HUB central antes de terminar em seu destino. Sua principal vantagem é a centralização de conexões em um dispositivo, que pode gerenciar todas as conexões. E a principal desvantagem é quando há algum problema no dispositivo central que geralmente causa uma parada em todo o sistema. Tipo de configuração mais comum atualmente.

Modelos De Referência

Antes da criação de um modelo de referência, quando uma empresa desejava desenvolver uma solução de computação em rede, ela precisava construir um sistema de informação baseado em tecnologias homogêneas o que significa que os dispositivos precisavam utilizar tecnologias padronizadas entre os diferentes hosts conectados a uma rede. Com a evolução da tecnologia e o aumento de dispositivos de um sistema de informação, houve a necessidade de se desenvolver um modelo que possibilitasse que diferentes tecnologias interoperassem dentro de uma rede de computadores. A International Organization for Standardization - ISO criou um modelo de camadas de protocolos chamado Open Systems Interconnection - OSI que se tornou modelo de referência para os sistemas de redes de computadores sobretudo para organização e interoperação dos protocolos de rede. Esse modelo foi originalmente organizado em sete camadas com funções bem definidas e posteriormente foi criado o modelo Transmission Control Protocol - TCP e Internet Protocol - IP que definiu um conjunto de protocolos de rede organizados em uma estrutura de quatro camadas que é o mais usado atualmente.

Camadas - Open Systems Interconnection

Camada de Aplicação (Mensagem) - Camada mais próxima do usuário em que ocorre a comunicação, responsável por operacionalizar os sistemas de informação definindo como ocorre a comunicação entre esses sistemas e os usuários e como as informações devem ser transmitidas e recebidas via protocolos existentes.

Camada de Apresentação (PPDU) - Responsável por definir a apresentação e a formatação dos dados. Essa camada tem por objetivo a compreensão dos dados considerando a sintaxe e a semântica das informações transmitidas pela rede, direcionando os dados para aplicações finais na camada de aplicação.

Camada de Sessão (SPDU) - Essa camada permite que os usuários, em diferentes hosts estabeleçam sessões de comunicação entre as aplicações.

Camada de Transporte (Segmentos) - Nessa camada, a função básica é aceitar dados da camada acima, dividi-los em unidades menores e determinar o tipo de serviço a ser executado com um protocolo orientado ou um não orientado à conexão.

Camada de Rede (Datagramas) (Pacotes) - Essa camada tem como objetivo controlar as operações da SUB-Rede, identificando e gerenciando a maneira como os pacotes de dados são roteados do host de origem até o host de destino e realizando também o endereçamento lógico dos hosts de rede.

Camada de Enlace (Quadros) - Essa camada tem como tarefa principal transformar um canal de comunicação em uma linha de dados livre de erros.

Camada Física (BITS) - Camada que trata a transmissão de sinais, o meio físico e onde está situada toda a parte de hardware de uma rede.

Camadas - Transmission Control Protocol

Encapsulamento - Essa técnica adiciona informações quando um dado é encaminhado de uma camada a outra. Operações de transporte de dados de forma controlada e com dados adicionais para controle é conhecida como encapsulamento.

Camada de Aplicação (HTTP e DNS) - Camada composta por protocolos de rede que são responsáveis pela operacionalização de sistemas e aplicações finais. Nessa camada, são definidos como os programas vão se comunicar com as aplicações e o gerenciamento das interfaces dessas aplicações.

Camada de Transporte (TCP e UDP) - Camada composta por protocolos de transporte de dados em rede que fornecem, à camada de aplicação, serviços de empacotamento e comunicação de duas formas, sendo uma delas via serviços orientados à conexão e a outra via serviços não orientados à conexão. Ela tem como função realizar e gerenciar conexões para garantir a integridade dos dados por meio de sequenciamento de pacotes segmentados no envio e recebimento de mensagens.

Camada de Internet (IP) - É responsável pela definição do endereçamento de um host de rede por meio do endereço de rede e também de roteamento dos pacotes de dados pelos diversos dispositivos de rede.

Camada Host de Rede - Como no modelo OSI, é a camada em que se localizam os dispositivos físicos da rede e as funções de enlace para acesso aos dispositivos físicos da rede. Entre suas atribuições, estão o monitoramento de tráfego de rede e o endereçamento em nível físico de dispositivos de rede para se realizar a transmissão de dados.

Protocolos De Redes

Protocolos Da Camada De Aplicação

Hyper Text Transfer Protocol (HTTP) - Protocolo usado em sistemas de WWW para representação de sistemas dentro de navegadores. Trata-se de um protocolo com intenso uso na atualidade, pois a grande parte dos sistemas da internet é executada com esse protocolo. O HTTP define como os clientes requisitam páginas aos servidores e como eles as transferem aos clientes. Esse protocolo está no coração da internet e é por meio desse padrão de comunicação em redes que as páginas de conteúdo são programadas e distribuídas via internet. WWW e HTTP são sistemas de padrões universalmente aceitos.

Domain Name System (DNS) - Protocolo utilizado para o sistema de nomes de domínio que faz a interconexão de URL, ou seja, nomes de endereços de sites da internet com endereços IP. Para que endereços de Internet tenham sua localização em um sistema distribuído mundialmente, é necessário que exista um sistema de nomes e domínios para a organização dos servidores na rede. Esse sistema é chamado Domain Name System (DNS) e seu objetivo é organizar os servidores para que endereços IP sejam convertidos em nomes.

Protocolos Da Camada De Transporte

User Datagram Protocol (UDP) - Protocolo de nível de transporte não orientado à conexão utilizado para transmissões que necessitam de maior velocidade, porém ele não garante a entrega dos dados. Exemplo: Streaming

Transmission Control Protocol (TCP) - Protocolo de nível de transporte orientado à conexão utilizado em aplicações que exigem que a sua totalidade e a sua integridade de dados sejam realizadas sempre com garantia de entrega. Exemplo: Internet Banking

Protocolos Da Camada De Internet

Internet Protocol (IP) - Protocolo mais conhecido de camada de Internet, uma vez que é o responsável pelo endereçamento lógico dos hosts de rede, informado pelo profissional de tecnologia da informação ou pelo próprio usuário para identificação única do host.

Protocolos Da Camada De Host De Rede

Carrier Sense Multiple Access With Collision Detection - Protocolo que utiliza um único meio de transmissão para suportar a transmissão de todos os hosts da rede. Seu funcionamento ocorre por meio de acesso múltiplo com detecção de onda portadora independentemente da topologia da rede. A transmissão é feita quando o cabo está livre e existe um controle de colisão quando mais de um host transmite dados ao mesmo tempo. Esse padrão é implementado nas redes padrão Ethernet.

Carrier Sense Multiple Access With Collision Avoidance - Protocolo que define o formato de utilização de um meio de comunicação compartilhado por meio de prevenção de colisão de onda portadora. Seu funcionamento é realizado a partir da análise do meio pelo qual o sinal será transmitido, e ao se verificar que o canal está livre, a transmissão é iniciada. Durante a transmissão em uma rede de computadores, cada host verifica o canal antes de transmitir e não transmitem quando percebem que o canal está ocupado.

Arquitetura e Tecnologia De Redes

Redes e SUB-Redes

Internet Protocol 4

Usado para fazer a identificação dos dispositivos em uma rede. Computadores e Servidores podem receber endereços IP. No caso dos Servidores na maior parte das vezes esse endereço é fixo. Não existem dispositivos com o mesmo número de IP. O IP é um número binário no caso do IPV4 e hexadecimal no caso do IPV6. Porém os usuários veem como numero decimal comum como 192.168.0.12. Cada conjunto de números antes dos pontos é chamado de Octeto.

Classes

Para saber a classe de um IP olhe para o primeiro Octeto.

Classe A - Endereços de IP entre 0 e 127

Classe B - Endereços de IP entre 128 e 191

Classe C - Endereços de IP entre 192 e 223

Classe D - Endereços de IP entre 224 e 239 - Reservados para Multicast

Classe E - Endereços de IP entre 240 a 255 - Reservados para Testes

Multicast - Quando se envia uma mensagem para um grupo específico dentro da rede.

Unicast - Quando se envia uma mensagem para apenas um endereço dentro da rede.

Broadcast - Quando se envia uma mensagem para todos os endereços existentes na rede.

Anycast - Quando se envia uma mensagem para o endereço que esteja mais próximo.

Máscaras

Classe A - 10.158.201.85

Observação - Tudo que está destacado pertence a REDE.

Observação - Tudo que não está em destaque pertence aos HOSTS.

Máscara - 255.0.0.0 - Rede.Host.Host.Host

Classe B - 172.16.189.85

Observação - Tudo que está destacado pertence a REDE.

Observação - Tudo que não está em destaque pertence aos HOSTS.

Máscara - 255.255.0.0 - Rede.Red.Host.Host

Classe C - 192.168.12.85

Observação - Tudo que está destacado pertence a REDE.

Observação - Tudo que não está em destaque pertence aos HOSTS.

Máscara - 255.255.255.0 = Rede.Red.Red.Host

Classe A - 255.0.0.0 - 256×256×256 Hosts

Classe B - 255.255.0.0 - 256×256 Hosts

Classe C - 255.255.255.0 - 256 Hosts

SUB-Redes

Uma SUB-Rede é quando uma rede é dividida em varias redes menores com o principal objetivo de evitar que elas se comuniquem Imagine que em uma faculdade a rede destinada aos alunos é a mesma destinada a secretaria. Como as redes não são separadas, qualquer aluno que tenha acesso a rede destinada aos alunos, terá acesso também a rede destinada a secretaria bem como os dados destinados a mesma. Por questões de segurança existe a necessidade de segmentar uma rede principal em algumas SUB-Redes. Agora, como calcular uma SUB-Rede?

Rede 192.168.1.0

Máscara De SUB-Rede - 255.255.255.192

Primeiro Passo - Encontrar o Octeto Misto = 255.255.255.192

Segundo Passo - Encontrar o Salto = $256 - 192 = 64$

Terceiro Passo - Determinar os Endereços de Rede

192.168.1.0

Salto 64 = $(0 + 64)$

192.168.1.64

Salto 64 = $(64 + 64)$

192.168.1.128

Salto 64 = $(128 + 64)$

192.168.1.192

Salto 64 = $(192 + 64)$

192.168.1.256

Quarto Passo - Determinar os Endereços de Broadcast

Rede 192.168.1.0

Sempre a Rede Seguinte (-1)

192.168.1.64 (-1)

Broadcast = 192.168.1.63

Rede 192.168.1.64

Sempre a Rede Seguinte (-1)

192.168.1.128 (-1)

Broadcast = 192.168.1.127

Rede 192.168.1.128

Sempre a Rede Seguinte (-1)

192.168.1.192 (-1)

Broadcast = 192.168.1.191

Rede 192.168.1.192

Sempre a Rede Seguinte (-1)

192.168.1.256 (-1)

Broadcast = 192.168.1.255

Quinto Passo - Determinar os Endereços de Host

Hosts da Rede 192.168.1.0

192.168.1.1 a 192.168.1.62

Broadcast 192.168.1.63

Hosts da Rede 192.168.1.64

192.168.1.65 a 192.168.1.126

Broadcast 192.168.1.127

Hosts da Rede 192.168.1.128

192.168.1.129 a 192.168.1.190

Broadcast 192.168.1.191

Hosts da Rede 192.168.1.192

192.168.1.193 a 192.168.1.254

Broadcast 192.168.1.255

Domain Name System - DNS

O IP é um número que faz a identificação de quem recebe ou envia dados em uma rede, porém como esse número considerado extenso para algumas pessoas, antes da criação dos servidores DNS, você teria que decorar o IP dos computadores para quem desejasse enviar algo. Agora imagine quantos números de telefone você tem decorado? Provavelmente é somente o seu e de mais algumas poucas pessoas, certo? Agora imagine a dificuldade em ter que decorar número de IP de um ou mais computadores espalhados pelo mundo. Para resolver esse problema surgiu o protocolo DNS, que é como se fosse uma agenda eletrônica. Basicamente funciona por meio de vários servidores distribuídos pelo mundo todo que são considerados a base da Internet. Esses servidores tem a função de identificar endereços de IP por meio de domínios e nomes. Quando você vai acessar algum site, você geralmente digita o nome do site ou sua URL e o navegador carrega o site automaticamente correto? Na verdade não, quando você digita o nome ou URL de um site basicamente o que você está fazendo é solicitar a um servidor de DNS que por meio do domínio em que ele está registrado, carregue o IP do mesmo e abra o site por meio de seu navegador.

Dynamic Host Configuration Protocol - DHCP

O DHCP permite atribuir endereços IP e outras informações de configuração a computadores clientes em uma rede local. Uma rede que possui um servidor DHCP disponível permite aos seus computadores obterem um endereço IP através de solicitação e atribuição automática pelo servidor. Um administrador de rede pode configurar um serviço para que determinado host receba o mesmo endereço IP toda vez que se conectar, ou um endereço IP temporário, diferente a cada conexão.

Ethernet e Internet Protocol 6

Ethernet

Ethernet é um protocolo de conexão que gerencia como os dispositivos e computadores se comunicam em uma Rede Local. Esse protocolo facilita o acesso e a troca de informações e arquivos entre máquinas que estão na mesma rede e também dá nome aos cabos de conexão presentes nos equipamentos.

Internet Protocol 6

O Internet Protocol 6 foi criado mediante a um aumento expressivo no número de computadores e outros dispositivos eletrônicos que se interconectavam às redes de computadores. Um endereço IPV4 permite uma quantidade de endereços para hosts de aproximadamente 4.3 Bilhões. Um endereço IPV6 é composto representa uma quantidade de endereços para hosts de 340 Undecilhões. Este número é suficiente para que possamos endereçar todos esses novos dispositivos na internet.

Gerenciamento De Redes

Padrões De Gerenciamento De Rede

SNMP (Simple Network Management Protocol) - Protocolo usado para realizar monitoramento de dispositivos e serviços de rede, que pode ser executado por dispositivos de diferentes arquiteturas e sistemas operacionais.

CMISE (Common Management Service Element) - A análise e a configuração de redes de computadores fazem uso de comandos em Prompt de Comando nos sistemas operacionais, para que pacotes de software sejam implementados no sistema e possam ser utilizados pelo administrador da rede para gerir os host e dispositivos de rede em geral.

Indicadores De Gerência

Latência - Especifica quanto tempo leva para os dados viajarem através da rede de um computador para outro, medida em Frações de Segundo.

Taxa de Transferência - Velocidade na qual os dados podem ser enviados através da rede e geralmente é medido em Bits por Segundo.

Perda De Pacotes - Pode ocorrer devido à capacidade de armazenamento de pacotes nos roteadores, considerando que estes possuem capacidade de memória limitada.

Confiabilidade

As redes de computadores são compostas por diversos equipamentos e todos esses dispositivos podem sofrer algum tipo de falha. A disponibilidade de uma rede é a capacidade que seus equipamentos possuem de se manterem em operação de forma ininterrupta dentro de um determinado período de tempo.

Mean Time Between Failures - Previsão por modelo estatístico e matemático do tempo médio entre as falhas da rede. Usado para prever as manutenções necessárias dentro de um sistema de redes de computadores.

Mean Time To Repair - Previsão por modelo estatístico e matemático do tempo médio para se realizar o reparo da rede após a ocorrência de uma falha.

Mean Time To Failure - Tempo de vida de uma rede que compreende os períodos alternados de operação de falhas. Este termo é utilizado para efetuar o cálculo de disponibilidade de uma rede de computadores por meio da função de frequência com que as falhas ocorrem e do tempo necessário para reparo.

Introdução Ao VLAN Trunk Protocol

Protocolo, desenvolvido pela Cisco e utilizado para configuração de uma Virtual Local Area Network, com o objetivo de facilitar a administração dos sistemas. Este protocolo define uma estrutura do tipo cliente e servidor, na qual as alterações são feitas necessariamente no servidor e replicadas aos clientes da rede. Esta técnica é utilizada por administradores de redes para melhorar o controle do sistema. Com o uso desse protocolo, o trabalho de configuração de redes locais virtuais é reduzido drasticamente, pois o gestor da rede configurará apenas um switch, que será o responsável pela função de distribuir e sincronizar as informações para os outros switches da rede. Esta tecnologia oferece redução de atividades de configuração e reconfiguração e minimização de erros através da centralização das configurações.

Sistemas Distribuídos

Exemplos De Sistemas Distribuídos

Ao abrir um e acessar uma página na internet, você está usando um sistema distribuído. Mas o que é um sistema distribuído? Um sistema distribuído é um conjunto de computadores que são interligados via rede, porém, para o usuário final das aplicações que são executadas através deles, aparenta ser um sistema único, como uma única máquina.

Compartilhamento De Recursos

Capacidade do sistema em compartilhar o acesso a quaisquer recursos utilizados por ele entre as máquinas que fazem parte da arquitetura. Apesar de não ser um objetivo exclusivo dos sistemas distribuídos, uma vez que também é um objetivo dos sistemas de rede, é uma característica muito importante. A vantagem de compartilhar recursos está na economia financeira, uma vez que, caso não haja tal possibilidade de compartilhamento, mais réplicas de um determinado recurso devem estar presentes em cada nó do sistema, o que impacta no custo. Como aspecto negativo associado a esse compartilhamento de recursos, temos a questão da segurança, uma vez que o fato de mais máquinas terem acesso ao recurso implica que o sistema possui mais pontos de acesso, que podem ser explorados.

Confiabilidade

A análise morfológica da palavra Confiabilidade nos mostra que ela se refere à probabilidade de um produto executar a sua função prevista de forma que atenda ou exceda às expectativas. Podemos confundir confiabilidade com algo relacionado à segurança do sistema, porém não tem relação alguma. A confiabilidade nos sistemas distribuídos é maior que nos sistemas centralizados, no entanto qualquer problema relacionado aos processos ou ao canal de comunicação pode surtir efeitos diretos sobre a execução do sistema.

Desempenho

Aumentar o desempenho de um sistema também é um objetivo dos sistemas distribuídos. Se fizermos uma comparação, os sistemas distribuídos, na maioria dos casos, apresentam melhor desempenho do que os sistemas centralizados. Isto ocorre porque, em um sistema distribuído, temos múltiplas instâncias, tanto de hardware quanto de software, para realizar o processamento necessário.

Middleware

Os sistemas distribuídos podem ser considerados como uma solução mais robusta em resposta aos sistemas puramente de rede, isso graças ao componente conhecido como Middleware. Ele é um dos fatores principais para o bom funcionamento de aplicações distribuídas. Middleware é uma camada oculta a qual se encontra entre os sistemas operacionais e os aplicativos que são executados neles. Portanto, é uma camada central, que permite o gerenciamento e a tradução de dados para interligar o sistema operacional com os programas, permitindo o funcionamento de aplicativos distribuídos.

Computação Em Cluster e Computação Em Grid

CLUSTER - *Esse tipo de computação é formado por um conjunto de máquinas com hardwares semelhantes, ou seja, máquinas que possuem características homogêneas. Quando falamos da parte de software da computação em cluster, temos algumas características importantes. Na maioria das vezes, o sistema operacional entre as máquinas é equivalente. Além disso, é frequente que um único programa funcione de forma subdividida em partes menores, e cada parte sendo executada em uma máquina, de forma distribuída, a fim de obter um aumento significativo de desempenho e menor tempo para executar determinada tarefa. Geralmente, as máquinas desse tipo de sistema são fortemente acopladas em suas ligações e, muitas vezes, podem até compartilhar a mesma memória RAM. Há sempre uma máquina que gerencia o funcionamento da aplicação entre todas as outras máquinas do sistema.*

GRID - *Esse tipo de computação é formado por um conjunto de máquinas com características diferentes, podendo o hardware e os sistemas operacionais serem de fabricantes diferentes. Com isso, temos uma característica heterogênea nessa computação onde essencialmente, um sistema de computação em Grid interliga vários Clusters.*

Diferenças Entre Clusters e Grids

Pode parecer que clusters e grids são a mesma coisa, mas existe uma característica fundamental que difere esses dois tipos de sistemas distribuídos. Para facilitar o entendimento das diferenças entre eles, podemos pensar que clusters são sistemas homogêneos, ou seja, são criados para executarem alguma tarefa específica que necessita de um alto poder de processamento e, portanto levaria muito tempo para ser executada em um computador convencional. Por sua vez, podemos pensar que os grids têm uma abordagem heterogênea, ou seja, são criados para executarem diferentes tarefas, de certa maneira relacionadas entre si, formando um centro de pesquisas de caráter Multidisciplinar. Uma maneira ainda mais simples de entender essa característica é pensar em um grid como um conjunto de dois ou mais clusters, cada um deles responsável por um certo tipo de pesquisa.

Computação Em Nuvem

O termo computação em nuvem se refere a uma tecnologia que possibilita acessar recursos e serviços via internet, sem a necessidade de instalações de softwares em seu computador. Dessa forma, é permitido que os usuários façam acessos por meio de qualquer dispositivo, seja ele um computador ou telefone celular.

Um exemplo de utilização de serviços em nuvem é quando você edita um documento utilizando o Google Docs e esse trabalho fica armazenado na nuvem.

O Papel Da Virtualização Em Sistemas Distribuídos

A vantagem da criação e utilização de redes virtuais advém do fato de que uma rede virtual específica para um determinado tipo de aplicação pode ser criada sobre uma rede física real, de forma que possa ser otimizada para aquela aplicação em particular, sem a necessidade de alterar as características da rede física. A virtualização de sistemas é uma alternativa interessante por permitir emular o hardware de uma máquina física, permitindo que várias máquinas virtuais, cada uma com um sistema operacional se comunicar. A principal vantagem da virtualização de sistemas está no fato de que aplicações já escritas e validadas, as quais dependem de um sistema operacional em específico e que necessitam se comunicar e interagir com outra aplicação em um sistema operacional diferente podem assim fazer através da virtualização dos sistemas operacionais, sem a necessidade de que a aplicação seja reescrita novamente ou recompilada.

Processos De Comunicação

Processos e Threads

Em sistemas operacionais, um conceito muito importante é o de Processo, o qual é definido como um programa em execução, e Threads que são fluxos ou linhas de execução dentro de um processo. Como exemplo de processo, podemos considerar um editor de texto executando em um desktop ou um aplicativo de calendário executado em um smartphone. Um processo pode possuir um ou mais threads permitindo que o programa execute mais de um trecho de código simultaneamente. O ato de gerenciar e executar múltiplas computações ao mesmo tempo é denominado concorrência. Por exemplo, uma máquina com processador contendo quatro núcleos de processamento executando dez programas diferentes necessita que o sistema operacional escalone fatias de tempo para que todos os programas consigam utilizar os núcleos de processamento por meio de revezamento. Estes programas estarão executando de forma concorrente, dando ao usuário a impressão de que cada programa está utilizando os recursos do computador de forma dedicada. Quando há execução de computações simultaneamente em diferentes núcleos de processamento, se tem o que é conhecido como paralelismo onde, processos e threads podem ser executados de forma concorrente e paralela.

Processos Cliente e Servidor

O modelo Cliente e servidor é o modelo de computação distribuída em que um programa é dividido na parte do Servidor e na parte do Cliente. O Servidor é executado em um processo, e o cliente, em outro processo, que pode estar na mesma máquina ou em uma máquina remota. O servidor tem acesso direto ao recurso de Hardware ou Software que o cliente quer usar. Geralmente um servidor atende às requisições de muitos clientes. O servidor é o mediador do acesso a grandes bases de dados, a um recurso de hardware caro ou a uma coleção importante de aplicações. Embora os processos cliente e servidor sejam separados em computadores diferentes, eles trabalham sempre em uma única aplicação. No contexto de comunicação entre processos, o processo que inicia a comunicação é o cliente, e o processo que espera pela requisição é o servidor.

Comunicação Entre Processos e Em Grupo

Após a criação dos processos, eles são executados de forma independente e possuem regiões de memória independentes. Entretanto, há aplicações que necessitam trocar dados entre processos. Para haver comunicação é necessária a utilização de mecanismos específicos. O sistema operacional implementa canais de comunicação entre processos em um mesmo computador ou em computadores diferentes. MPI é um padrão de comunicação que permite a troca de mensagens entre processos pertencentes a um grupo de processos criado no início da execução do programa. Cada processo recebe um identificador único dentro deste grupo, facilitando a comunicação e a identificação sobre qual é o processo emissor e o receptor da mensagem. Dentro de um mesmo processo, é possível criar threads que compartilham memória através do acesso à memória compartilhada.

Conceito De Sockets

A padronização da interface da camada de transporte permitiu que os protocolos de troca de mensagens fossem utilizados pelos programadores no desenvolvimento de aplicações que envolvam mais de uma máquina. Uma mensagem enviada de um processo para outro deve passar pela rede, e uma forma de realizar essa comunicação é utilizando uma interface denominada Socket que é uma interface entre a camada de aplicação e o protocolo da camada de transporte. O Socket é um terminal de comunicação para o qual uma aplicação pode escrever dados que devem ser enviados pela rede e do qual pode ler os dados que chegam.

Aspectos De Projeto Dos Sistemas Distribuídos

Abertura

No contexto de sistemas distribuídos, se refere a quanto é fácil integrar e alterar tecnologias sem que o sistema seja comprometido. Importante observar que independentemente do que o termo abertura possa passar, é uma coisa positiva em sistemas distribuídos. É importante ter em mente esse conceito, pois à primeira vista, um sistema mais aberto parece ser uma coisa negativa no sentido de estar mais vulnerável a falhas, mas isso não é verdade, pois a segurança não tem relação nenhuma com a abertura.

Concorrência

Capacidade do sistema em poder ser acessado e utilizado de maneira simultânea por vários usuários. Aqui, cabe novamente a ressalva de que, no contexto de sistemas distribuídos, o termo concorrência não tem uma conotação negativa como ocorre, por exemplo, no comércio. Apenas se refere a um sistema que dá suporte a acessos simultâneos.

Escalabilidade

Escalabilidade é um termo comum em redes de computadores e está relacionado à capacidade de o sistema poder ser ampliado ou reduzido para suportar, por exemplo, uma maior quantidade de acessos simultâneos ou realizar uma tarefa mais rapidamente. 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 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 conhecido do mercado, digamos, com dez nós. Apesar de a ideia ser excelente, você nota que a quantidade de usuários que utiliza sua plataforma cai drasticamente entre os meses de novembro e junho. Supondo que você paga para esse provedor 150 reais mensais, para que este disponibilize os dez nós de maneira contínua. Não seria interessante que, nos meses de menor demanda, você diminuísse a quantidade de servidores pela metade pagando a quantia de 75 reais mensais nesse período? Nesse cenário, sua economia seria de 600 reais que você poderia investir em outros projetos. Esse é um exemplo típico de escalabilidade que diminui a capacidade do sistema.

Escalonamento Horizontal - Permite o aumento de máquinas. Por exemplo, aumento de servidores em uma determinada empresa, fazendo com que assim haja uma divisão da carga de trabalho, não sobrecarregando um servidor específico.

Escalonamento Vertical - Permite aumentar recursos de processamento, como CPU e memória em máquinas existentes. Quando se deseja maior desempenho.

Heterogeneidade

Quando temos heterogeneidade, estamos falando de um sistema que contenha em sua composição máquinas de sistemas operacionais, recursos de hardware e até mesmo fabricantes diferentes. Esse é um dos aspectos mais frequentes de um sistema distribuído. Geralmente, o sistema é composto por máquinas de diversas características diferentes que se comunicam para manter o funcionamento.

Segurança

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 que replicam tal aplicação e a comunicação entre essas máquinas sempre ocorre por meio de redes de comunicação cabeadas. Importante sempre considerar aspectos de segurança no projeto de sistemas distribuídos.

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 abertas para que clientes possam acessar tais serviços, dessa forma, um Hacker pode enviar mensagem a qualquer uma delas.

Redes De Computadores Não São Seguras - Remetentes de mensagens podem ser alterados e endereços IP podem estar duplicados, de forma que alguém malicioso possa receber as mesmas mensagens de um destinatário válido.

A Validade Das Chaves Criptográficas Deve Ser Limitada - Quanto mais tempo uma mesma chave estiver ativa, maiores são as chances de estar comprometida.

Algoritmos De Criptografia Podem Ter Falhas - Atualmente, a melhor prática é divulgar publicamente os algoritmos de criptografia para que a comunidade e as entidades especializadas possam validá-lo 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 diversos 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 que tentam descobrir senhas por tentativa e erro.

Tolerância a Falhas

Capacidade que o sistema distribuído tem para autor recuperar na ocorrência de uma ou mais falhas. Sua ideia básica é de que os processos da nossa aplicação sejam replicados em grupos, isso faz com que o sistema tenha uma proteção contra falhas relacionadas a processos. Para conseguir criar projetos resilientes, 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 verificando qual grupo de processos que a nossa aplicação deverá conter. Organizamos em grupos os processos que são considerados idênticos. Quando uma mensagem é enviada a um grupo de nosso sistema, a ideia é que todos os processos membros desse grupo também recebam a mensagem. Se ocorrer uma falha em um processo, outro processo desse grupo deve tratar a mensagem.

Grupos Simples - Todos os processos são iguais e todas as 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. A desvantagem que encontramos nesse tipo de grupo é que a tomada de decisão tende a demorar mais, porque cada decisão deve ser priorizada pelos processos, tendo uma votação antes da tomada de decisão.

Grupos Hierárquicos - Existem alguns 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 executar. A grande vantagem desse tipo de grupo é que as decisões são centralizadas, portanto temos mais agilidade na tomada de decisão. Já a grande desvantagem apontada é que, caso ocorra uma falha no processo coordenador, o serviço todo para.

Transparência

Olhando da perspectiva de quem usa o sistema, se refere a quanto um usuário desconhece sobre o funcionamento interno do sistema, e isso é considerado uma característica positiva em um sistema distribuído.

Virtualização e Containerização

Arquitetura De Virtualização

A maioria das pessoas quando ouve falar em máquinas virtuais pensa em um sistema operacional dentro de um software como o VirtualBox da Oracle, instalado em uma máquina física com um sistema operacional instalado. Partindo desse contexto, podemos destacar algumas características. Esse sistema operacional instalado dentro da máquina física se refere a uma máquina virtual, pois é similar a máquina física, porém puramente emulada via software. O software que permite emular uma máquina física é chamado de Hypervisor e é responsável por desacoplar a máquina física da virtual, bem como alocar os recursos da máquina física de acordo com a necessidade da máquina virtual.

Conceito De Containerização

Uma das tecnologias mais populares que temos atualmente é o uso de contêineres para a execução de sistemas dos mais variados tipos. Isso ocorre devido à facilidade e à flexibilidade que advêm do uso dos mesmos. O contêiner funciona como uma tecnologia que dá o suporte para o funcionamento de uma aplicação e pode ser considerado a emulação de nossa aplicação. Quando a aplicação é executada através de um contêiner, ela tem todas as bibliotecas e os elementos necessários para o funcionamento disponíveis dentro do contêiner. Uma maneira simples para entender o que são os chamados contêineres é imaginar que eles permitem a criação de ambientes virtuais isolados e independentes para serem utilizados por aplicações, similar ao resultado do uso de máquinas virtuais. Entretanto, um grande diferencial está no fato de que os contêineres são mais leves que as máquinas virtuais, por possuírem uma arquitetura mais otimizada. Outra grande vantagem que não há, obrigatoriamente, a necessidade de instalar um sistema operacional completo, visto que as plataformas de containerização aproveitam bibliotecas compartilhadas com o sistema operacional hospedeiro. Por essa razão, os contêineres ocupam menos espaço em disco e consomem menos RAM e processamento que as máquinas virtuais. Existem várias implementações para contêineres, o mais popular entre eles é o Docker.

O Papel Da Containerização

Os sistemas distribuídos fazem uso dos contêineres no contexto de micro serviços. A ideia dos micro serviços está associada a empresas que possuem sistemas altamente dinâmicos e ao termo modularidade. Caso o sistema possua uma arquitetura monolítica que tem por característica uma forte dependência entre elementos, será muito difícil substituir alguns desses elementos sem causar uma interrupção completa no sistema. Em uma arquitetura baseada em micro serviços esses componentes têm um baixo grau de dependência, de forma que, caso você tenha de fazer uma substituição, terá um impacto bem menor.