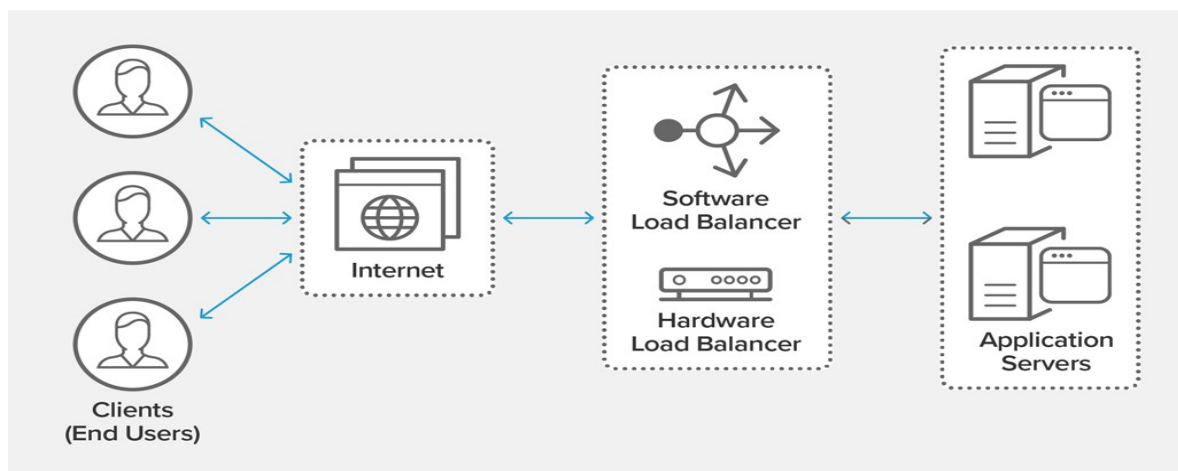


Sistemas Distribuídos – Manoel Campos

UANDERSON FERREIRA DE OLIVEIRA

Escalabilidade de Sistemas Distribuídos



Clients(End Users)

São todos aqueles que utilizam algum serviço disponibilizado pela Internet, através de um DNS valido na web, gerando diariamente um grande volume de requisições aos provedores de serviços.

Internet

E o meio/ a ponte que conecta o Client/Usuário final as aplicações web que estão sendo processadas nos servidores, ou seja, e o canal de toda a comunicação que ocorre entre clientes e servidores, por meio dos diversos protocolos da web.

Na imagem acima, a internet recebe as requisições dos usuários e envia para um balanceador de cargas que irá cuidar de realizar o gerenciamento das requisições aos servidores para garantir eficiência e um melhor tempo de resposta para o usuário, pois quando um usuário envia uma requisição ele necessita de uma resposta que normalmente é um conteúdo que deseja, podendo estar em diversos formatos aceitos pela web.

Software Load Balancer e Hardware Load Balancer

O balanceamento de carga refere-se à distribuição eficiente do tráfego de rede de entrada em um grupo de servidores de back-end, também conhecido como *uma fazenda(farm)* de servidores ou *uma piscina(pool)* de servidores.

Um balanceador de cargas atua como o “traffic cop” sentado na frente de seus servidores e encaminhando solicitações de clientes em todos os servidores capazes de atender a essas solicitações de maneira que maximize a velocidade e a utilização da capacidade e garanta que nenhum servidor esteja sobrecarregado, o que pode degradar o desempenho. Se um único servidor cair, o balanceador de carga redireciona o tráfego para os servidores online restantes. Quando um novo servidor é adicionado ao grupo de servidor, o balanceador de carga automaticamente começa a enviar solicitações para ele.

A diferença principal entre Software Load Balancer e Hardware Load Balancer é que enquanto um utiliza sistemas para a realização do balanceamento como o nginx, o outro se faz através de equipamentos físicos como por exemplo: roteadores e switches que basicamente podem ser vistos como um dispositivo que conecta e envia informações a todos os elementos da sua rede.

Application Servers

São servidores que disponibiliza um ambiente para a instalação e execução de aplicações diversas, centralizando e dispensando a instalação em computadores clientes. Os servidores de aplicação também são conhecidos por middleware. E onde as aplicações estão sendo processadas e também e onde ficam armazenados os arquivos/conteúdo de um determinado serviço.

Como a escalabilidade de tamanho é alcançada em um arquitetura destas e como o uso de recursos pode ser otimizado para reduzir consumo de energia ?

Através do escalonamento Horizontal (Horizontal Scaling ou Up Scaling), que consiste no alocamento de novos servidores para que seja possível atender a demanda de forma eficiente em conjunto com o balanceamento de cargas que será aplicado a estrutura do sistema, vale ressaltar que tal processo deve ser automatizado, afim de evitar possíveis perdas de desempenho computacional. Hoje em dia, a computação em nuvem possibilita uma automatização mais rápido, além de dispor de diversas opções de provedores Cloud como Amazon(AWS), Google e Microsoft(Azure), entre outras que usufruem de recursos como a virtualização para fornecer ambientes virtuais conhecidos como máquinas virtuais(VM) ou contêineres, gerando um ganho relevante no compartilhamento de recursos entre si.

Além disso, se têm uma redução no consumo de energia, pois é feito uma realocação de recursos, ou seja, um migração para nuvem onde reduz o uso de máquinas físicas e posteriormente o consumo de energia, que em muitos casos pode gerar grandes custos para uma empresa.

Como garantir que os recursos computacionais sejam alocados de acordo com a demanda ?

Através de uma implementação de balanceamento de cargas eficiente juntamente com recursos virtualizados e automatizados, que trabalharam em conjunto conforme a demanda aumenta e diminui, gerenciando os recursos conforme a necessidade e provendo novos recursos quando necessários.

Referências Bibliográfica:

<https://www.nginx.com/resources/glossary/load-balancing/>

https://docs.google.com/presentation/d/e/2PACX-1vQr5t9iBtomeqso65zXuYdo0gkWbd9rdH4olcu14V-In330CViYvVmU3qC_P0OTs-jvB0HdE5bLPmVw/pub?start=false&loop=false&delayms=60000#slide=id.p58