

Balanceamento de carga

Origem: Wikipédia, a enciclopédia livre.

Todo o hardware tem o seu limite, e muitas vezes o mesmo serviço tem que ser repartido por várias máquinas, sob pena de se tornar congestionado. Estas soluções podem-se especializar em pequenos grupos sobre os quais se faz um balanceamento de carga: utilização do CPU, de armazenamento, ou de rede. Qualquer uma delas introduz o conceito de clustering, ou *server farm*, já que o balanceamento será, provavelmente, feito para vários servidores.

Em rede de computadores, o **balanceamento de carga** é uma técnica para distribuir a carga de trabalho uniformemente entre dois ou mais computadores, enlaces de rede, UCPs, discos rígidos ou outros recursos, a fim de otimizar a utilização de recursos, maximizar o desempenho, minimizar o tempo de resposta e evitar sobrecarga. Utilizando múltiplos componentes com o balanceamento de carga, em vez de um único componente, pode aumentar a confiabilidade através da redundância.

Índice

Balanceamento de armazenamento (*storage*)

- Soluções
- Load Balancer em Cloud

Balanceamento de rede

- Exemplo

Balanceamento de CPU

Ver também

- Ligações externas

Balanceamento de armazenamento (*storage*)

O balanceamento do suporte de armazenamento permite distribuir o acesso a sistemas de ficheiros por vários discos (software/hardware RAID), pelo que derivam ganhos óbvios em tempos acesso. Estas soluções podem ser dedicadas ou existir em cada um dos servidores do cluster.

Soluções

- RAID
- Storage Area Network*
- Load Balancer - Cloud Computing oferecido como IaaS

Load Balancer em Cloud

É oferecido como serviço, não requer nenhum hardware, pois o serviço é prestado remotamente. Todo o tráfego é direcionado para um IP do provedor contratado e este é distribuído através da internet ou de rede local para os servidores de destino. Os IPs de destino podem estar tanto no datacenter do provedor quanto em outros datacenter. O serviço geralmente é cobrado por quantidade de requisições simultâneas e por GB de dados transferidos.

Balanceamento de rede

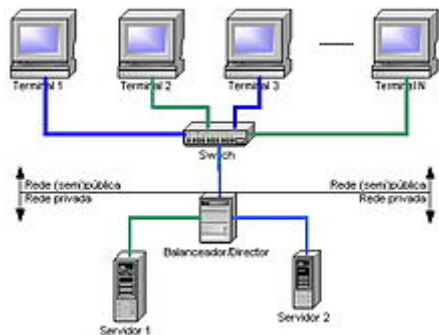


Figura 1 - Balanceamento de carga (NAT).

O balanceamento da utilização da rede passa sobretudo por reencaminhar o tráfego por caminhos alternativos a fim de descongestionar os acessos aos servidores. Este balanceamento pode ocorrer a qualquer nível da camada OSI.

- Equal-cost multi-path routing
- Shortest Path Bridging (IEEE 802.1aq)

A **Figura 1** sugere a existência de um mecanismo/dispositivo responsável pelo balanceamento (**director**). Na verdade, ele pode existir sob várias formas, dependendo do(s) serviço(s) que se pretende balancear. Este director serve também de interface entre o cluster de servidores e os clientes do(s) serviço(s) - tudo o que os clientes conhecem é o endereço semi-público deste servidor. Esta abordagem (clássica) é algo limitada, em termos de escalabilidade, ao número de tramas que o director consegue redireccionar, principalmente devido à velocidade dos buses das placas de rede. Existem, no entanto, outras soluções mais complexas que tiram melhor partido das características do protocolo TCP/IP em conjunto com *routing* especializado (**NAT**, **IP Tunneling**, **Direct Routing**).

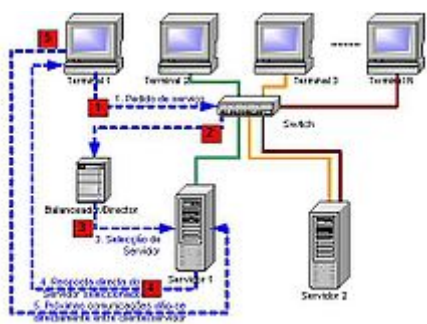


Figura 2 - Balanceamento de carga (*Direct Routing*).

Exemplo

Um exemplo de um site a utilizar técnicas de balanceamento de carga é a própria Wikimedia Foundation e os seus projectos. Em Junho de 2004, a carga era balanceada usando uma combinação de:

- Round robin DNS, que distribui os pedidos uniformemente para um dos três servidores de cache Squid;
- Estes servidores de *cache* usam os tempos de resposta para distribuir os pedidos para cada um dos sete servidores de páginas. Em média, os servidores Squid já têm em *cache* páginas suficientes para satisfazer 75% dos pedidos sem sequer consultar os servidores de páginas;
- Os scripts PHP que formam a aplicação distribuem a carga para um de vários servidores de base de dados dependendo do tipo do pedido, com as atualizações indo para um servidor primário e as consultas para um ou mais servidores secundários.

Balanceamento de CPU

Este tipo de balanceamento é efetuado pelos sistemas de processamento distribuído e consiste, basicamente, em dividir a carga total de processamento pelos vários processadores no sistema (sejam eles locais ou remotos).

Ver também

- Sistema de processamento distribuído
- Sistemas distribuídos

Ligações externas

- Linux High Availability Project (<http://www.linux-ha.org/>)
- NAT based SLB (<http://www.oreilly.com/catalog/serverload/chapter/ch07.html>)
- Load Balancer na base OpenBSD relayd (<http://www.openbsd.org/cgi-bin/man.cgi?query=relayd&apropos=0&sektion=8&manpath=OpenBSD+Current&arch=i386&format=html>)

Obtida de "https://pt.wikipedia.org/w/index.php?title=Balanceamento_de_carga&oldid=51038396"

Esta página foi editada pela última vez às 23h32min de 16 de janeiro de 2018.

Este texto é disponibilizado nos termos da licença Atribuição-CompartilhaIgual 3.0 Não Adaptada (CC BY-SA 3.0) da Creative Commons; pode estar sujeito a condições adicionais. Para mais detalhes, consulte as condições de utilização.