

Arquitetura e Organização de Computadores II

Computadores em escala warehouse

Prof. Nilton Luiz Queiroz Jr.

Escala Warehouse

- Computadores em escala Warehouse (Warehouse Scale Computer - WSC) são a base dos serviços de internet;
 - Buscas;
 - Redes sociais;
 - Compartilhamento de vídeo;
 - Serviços de e-mail;
- Tem arquitetura e operação diferente da de data centers;
- Agem como grandes máquinas e tem alto custo;
 - Cerca de 150 milhões, incluindo:
 - Instalações;
 - Infraestrutura de refrigeração;
 - Equipamentos de rede;

Escala Warehouse

- WSC podem ser vistos como descendentes de supercomputadores em questões como:
 - Inovação;
 - Distribuição de energia;
 - Monitoramento de operação e refrigeração;
- Supercomputadores são usados com maior finalidade de cálculos para cientistas e engenheiros;
- WSC são usados para fornecer ao mundo Tecnologia da informação;
 - Atendem muito mais usuários que supercomputadores;



Escala Warehouse

- Algumas **características** dos computadores em escala warehouse são:
 - Custo-desempenho;
 - Como os **custos são muito altos**, é necessário se reduzir na medida do possível os gastos com u WSC
 - Eficiência energética;
 - O trabalho realizado por joule em um WSC é muito importante;
 - Pode inclusive influenciar no sistema de refrigeração;
 - **Confiabilidade por meio de redundância**;
 - Os sistemas devem estar disponíveis por muito tempo;
 - Redundância auxilia a mascarar falhas;
 - Boa **interface de E/S com rede**;
 - Rede boa o suficiente para manter os dados coerentes entre os WSC e comunicação externa;

Escala Warehouse

- Cargas de trabalho interativas e em lote;
 - Além da interação com os usuários, warehouse devem ter aplicações executando em lote para cálculo de metadados úteis para serviços de interação;
 - Buscas tem interface com usuário e precisam “vasculhar” os endereços e organizá-los, as vezes até de acordo com o perfil do usuário;
- Amplo paralelismo;
 - Como existem muitos dados para processar, as aplicações em lote podem processar de maneira independente;
 - Paralelismo em nível de dados
 - Aplicações que fazem uso de interação com o usuário atendem diversos usuários ao mesmo tempo;
 - Paralelismo em nível de requisição
 - Leituras e escritas quase nunca são dependentes;



Escala Warehouse

- Problemas e oportunidades com escalabilidade
 - Redução de custos ao “Iniciar” um warehouse;
 - Como a quantidade de computadores personalizados é muito alta, é possível reduzir custos em aquisição por volume;
 - Quanto mais computadores maior a quantidade de falhas;



WSC x Cluster

- Os precursores dos computadores WSC são os clusters de computadores;
- São mais focados em aplicações científicas e de alto desempenho;
- Em geral clusters tem processadores e redes mais rápidas que os WSC;
- Clusters enfatizam paralelismo em nível de thread e de dados;
 - WSCs enfatizam paralelismo em nível de requisição;



WSC x datacenter

- Data centers convencionais coletam máquinas e software de terceiros de muitas partes de uma organizam e executam centralmente para outros;
- Muitos serviços em uma quantidade menor de máquinas isoladas;
 - Comum uso de máquinas virtuais;
- São mais heterogêneos que os WSC;
 - Tanto em Hardware quanto em software;
- Em WSCs é mais comum que os softwares sejam desenvolvidos pelos próprios programadores;
 - Tanto pela extensão de um software ja existente quanto pela criação desde o início de um software;
- Datacenters não tem a escala de um WSC;

Modelo de programação em WSC

- Além dos serviços como busca, compartilhamento de vídeo e redes sociais, computadores em escala warehouse executam aplicações em lote;
 - Conversão de vídeos para outros formatos;
 - Criação de índices de busca;
- Para isso existem alguns frameworks para auxílio a programação em warehouses:
 - MapReduce;
 - Hadoop;
 - Uma versão open source do MapReduce;



Map Reduce

- Inspiradas nas funções de Lisp
 - Map;
 - Aplica uma função fornecida pelo programador a cada registro da entrada;
 - Executado por milhares de computadores, cada um produzindo um resultado de pares chave e valor;
 - Reduce;
 - Coleta resultados de tarefas distribuídas e as junta de acordo com uma função pré definida pelo programador;



Modelo de programação em WSC

- Para alta disponibilidade usam replicação de dados entre diferentes servidores;
- Além disso, WSCs usam consistência relaxada;
 - É importante que multipla replicações concordem eventualmente;
 - Não é necessário que as réplicas estejam de acordo o tempo todo;
 - Se torna mais fácil escalar sistemas com esse tipo de consistencia;

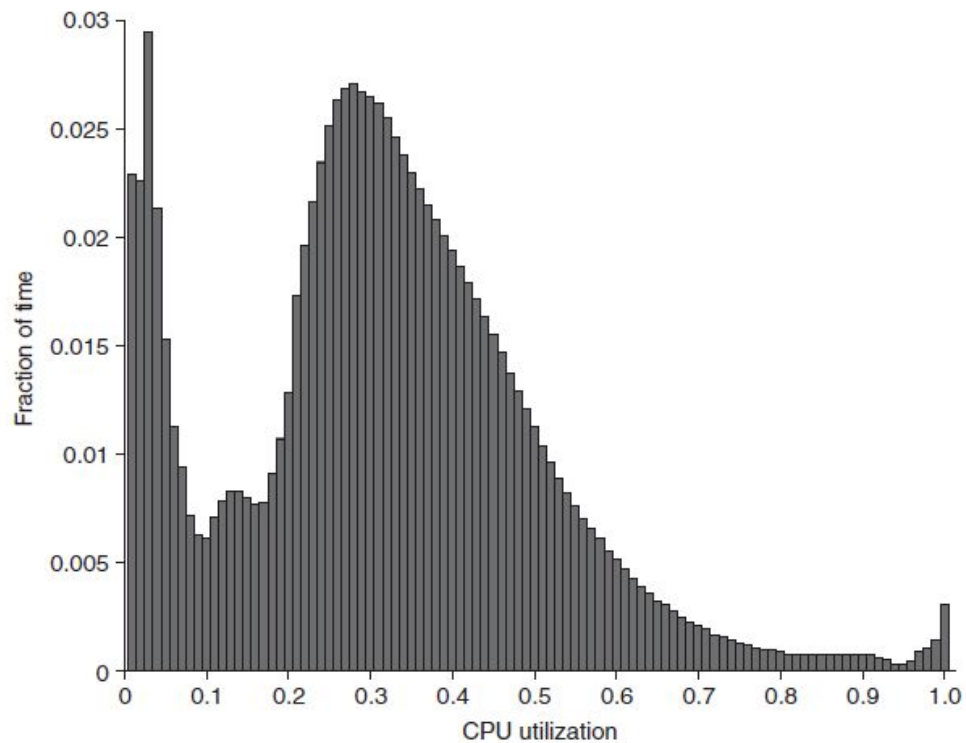


Demanda para WSC

- A demanda de processamento varia muito e de maneira considerável;
 - Eventos “sociais” podem alterar a demanda por processamento;
 - Véspera de Natal aumenta a quantidade de vendas;
 - Compartilhamento de fotos é maior após o Halloween;
 - Dados coletados em 5000 servidores do google ao longo de 6 meses mostram que:
 - A utilização média em servidores é de 100% em menos de 0,5% dos servidores;
 - A maior parte dos servidores opera entre 10 - 50% de utilização média;



Demanda para WSC



Arquitetura de um computador WSC

- WSCs são interconectados por redes, e fazem uso de hierarquia em redes para a conexão;
- Os servidores são distribuídos em racks switches;
 - Em geral os racks têm espaço para 48 servidores;
 - Switches de 48 servidores;
 - Esses racks oferecem de 2 a 8 uplinks;
 - Uplinks são usados para ir para um nível acima na hierarquia da rede;
 - Larguras de banda de 6 a 24 vezes menor (oversubscription);
 - 48/2;
 - 48/8;
 - Send e recieves para diferentes racks causam impacto na performance;

Armazenamento WSC

- O armazenamento em WSC pode ser feita de duas principais maneiras:
 - Pode-se colocar armazenamento dentro dos próprios servidores;
 - Menor performance;
 - Menor custo;
 - Armazenamento ligado a rede (NAS - Network Attached Storage);
 - Mais custoso;
 - Maior desempenho;
 - Técnicas RAID para melhorar a confiabilidade do armazenamento;
- A alternativa mais comum em WSC são os discos locais aos servidores;
 - Softwares de armazenamento lidam com confiabilidade e conectividade;
 - Google File System (GFS): usa discos locais e 3 replicas;

Switch de Arrays

- Existem arrays de switch, que interconectam um grupo de racks;
- Tem mais banda de bissecção que um switch rack;
 - Reduz problemas como oversubscription;
 - Custo maior devido a largura de banda de um switch crescer em ordem quadrática;



Hierarquia de memória em WSC

- Servidores podem acessar DRAM e discos em outros servidores usando uma interface estilo NUMA, tendo limitações:
 - Locais: o processador para acessar seu proprio hardware;
 - Entre rack switch;
 - Entre switch de array;

	Local	Rack	Array
DRAM latency (microseconds)	0.1	100	300
Disk latency (microseconds)	10,000	11,000	12,000
DRAM bandwidth (MB/sec)	20,000	100	10
Disk bandwidth (MB/sec)	200	100	10
DRAM capacity (GB)	16	1040	31,200
Disk capacity (GB)	2000	160,000	4,800,000

Hierarquia de Memória

- A maior parte das aplicações cabe em um único array dentro de um WSC;
- Aplicações que precisam de mais de um array podem particionar os dados entre arrays diferentes;
- As operações sobre o conjunto de dados são enviadas para os servidores que armazenam as diferentes partes e os resultados são reunidos pelo computador cliente;



Referências

HENNESSY, John L.;PATTERSON, David A. Computer architecture: a quantitative approach. Elsevier, 2011.

