



ECM 245

Arquitetura e
Organização de
Computadores

5ª-feira

07h40-09h20

Sala: H204

AULA 16

Processamento paralelo: Clusterização.

ECM 245

Arquitetura e
Organização de
Computadores

5ª-feira

07h40-09h20

Sala: H204

AULA 16

Introdução

Como já visto anteriormente, desde da última década observa-se um crescente número de aplicações que exigem cada vez mais “poder” computacional.

A demanda por poder computacional cresce em uma velocidade maior que os sistemas tradicionais conseguem acompanhar.

Uma alternativa para suprir esta necessidade é a implantação de sistemas “fracamente acoplados”, como cluster de computadores.



ECM 245

Arquitetura e
Organização de
Computadores

5ª-feira

07h40-09h20

Sala: H204

AULA 16

Definição

Como já verificado na aula anterior, o Cluster (aglomerado) é um sistema interligando no mínimo dois computadores, cujo objetivo é fazer com que todo o processamento da aplicação seja distribuído entre os processadores, de uma forma transparente, ou seja, aparentando um único processador.



ECM 245

Arquitetura e
Organização de
Computadores

5ª-feira

07h40-09h20

Sala: H204

AULA 16

Quando começou a utilização?

A tecnologia Cluster é mais antiga do que se imagina. Ela começou com as máquinas de alto desempenho (supercomputadores). O desenvolvimento dessa tecnologia para PC's foi motivada devido ao alto custo dos supercomputadores, começando em 1994 com o desenvolvimento do cluster *Beowulf* pela NASA.



ECM 245

Arquitetura e
Organização de
Computadores

5ª-feira

07h40-09h20

Sala: H204

AULA 16

Configuração Cluster

Um Cluster pode possuir vários tipos de configuração diferentes, seja com hardware ou software. Os tipos mais comuns de Clusters Linux são:

Cluster de processamento paralelo

Cluster de disponibilidade

ECM 245

Arquitetura e
Organização de
Computadores

5ª-feira

07h40-09h20

Sala: H204

AULA 16

Cluster de processamento paralelo

Cada vez que o Cluster recebe uma tarefa para executar (por exemplo, a renderização de um filme), já previamente preparada para “rodar” em processamento paralelo, o Cluster divide os “pedaços” da tarefa que cada uma das máquinas deve realizar. Um Cluster paralelo é ideal para executar poucas tarefas, mas que exigem grande quantidade de processamento.



ECM 245

Arquitetura e
Organização de
Computadores

5ª-feira

07h40-09h20

Sala: H204

AULA 16

Cluster de disponibilidade

Esse tipo de Cluster funciona como um gerenciador de tarefas, ou seja, cada máquina trabalha sozinha, porém a distribuição das tarefas é realizada de tal forma que os processadores estejam sempre trabalhando na capacidade máxima. Nesse tipo de Cluster é vital a implementação de um **sistema de filas** com vários níveis de prioridade diferentes. Esse tipo de cluster é ideal para trabalhar com grande quantidades de tarefas que necessitam de pequenas ou médias capacidades de processamento.

ECM 245

Arquitetura e
Organização de
Computadores

5ª-feira

07h40-09h20

Sala: H204

AULA 16

Beowulf

O primeiro Cluster Beowulf foi criado em 1994 na CESDIS, uma subsidiária da NASA. Ele era formado por 16 PCs 486 DX-100 ligados em rede. Para manter a independência do sistema e baixar os custos, os desenvolvedores optaram por utilizar o Linux. Diferente de outros tipos de Cluster, ele não exige uma arquitetura específica tão pouco máquinas homogêneas.



ECM 245

Arquitetura e
Organização de
Computadores

5ª-feira

07h40-09h20

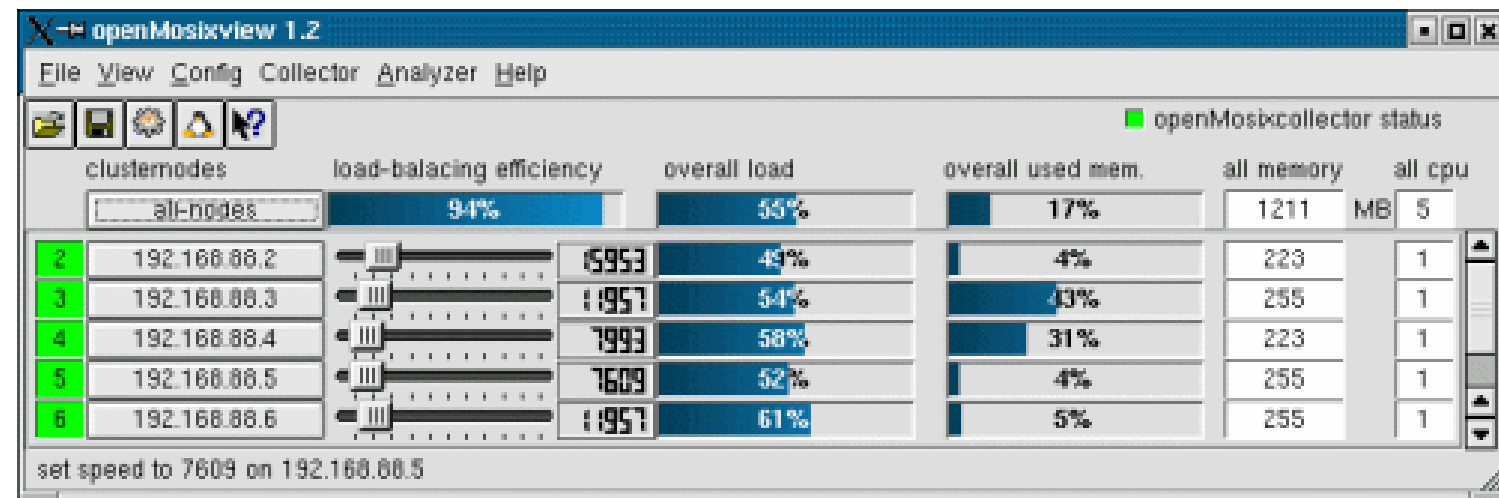
Sala: H204

AULA 16

OpenMosix



Este Cluster trabalha com distribuição de processos, que ao detectar o alto volume de processamento, migram as instâncias entre as máquinas do cluster, sendo processadas simultaneamente, sem a necessidade de adequação do código. (princípio da elasticidade)





ECM 245

Arquitetura e
Organização de
Computadores

5ª-feira

07h40-09h20

Sala: H204

AULA 16

| Rank | System | Cores | Rmax (TFlop/s) | Rpeak (TFlop/s) | Power (kW) |
|------|---|------------|-------------------|--------------------|---------------|
| 1 | Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/SC/Oak Ridge National Laboratory United States | 2,414,592 | 148,600.0 | 200,794.9 | 10,096 |
| 2 | Sierra - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States | 1,572,480 | 94,640.0 | 125,712.0 | 7,438 |
| 3 | Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China | 10,649,600 | 93,014.6 | 125,435.9 | 15,371 |
| 4 | Tianhe-2A - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000 , NUDT National Super Computer Center in Guangzhou China | 4,981,760 | 61,444.5 | 100,678.7 | 18,482 |
| 5 | Frontera - Dell C6420, Xeon Platinum 8280 28C 2.7GHz, Mellanox InfiniBand HDR , Dell EMC Texas Advanced Computing Center/Univ. of Texas United States | 448,448 | 23,516.4 | 38,745.9 | |
| 6 | Piz Daint - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland | 387,872 | 21,230.0 | 27,154.3 | 2,384 |
| 7 | Trinity - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/NNSA/LANL/SNL United States | 979,072 | 20,158.7 | 41,461.2 | 7,578 |
| 8 | AI Bridging Cloud Infrastructure (ABCI) - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan | 391,68 | 19,880.0 | 32,576.6 | 1,649 |
| 9 | SuperMUC-NG - ThinkSystem SD650, Xeon Platinum 8174 24C 3.1GHz, Intel Omni-Path , Lenovo Leibniz Rechenzentrum Germany | 305,856 | 19,476.6 | 26,873.9 | |
| 10 | Lassen - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, Dual-rail Mellanox EDR Infiniband, NVIDIA Tesla V100 , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States | 288,288 | 18,200.0 | 23,047.2 | |

Exemplos Reais

<https://www.top500.org/lists/2019/06/>



<https://www.top500.org/>



ECM 245

Arquitetura e
Organização de
Computadores

5ª-feira

07h40-09h20

Sala: H204

AULA 16

Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband

| | |
|----------------------------|---|
| Site: | DOE/SC/Oak Ridge National Laboratory |
| System URL: | http://www.olcf.ornl.gov/olcf-resources/compute-systems/summit/ |
| Manufacturer: | IBM |
| Cores: | 2,414,592 |
| Memory: | 2,801,664 GB |
| Processor: | IBM POWER9 22C 3.07GHz |
| Interconnect: | Dual-rail Mellanox EDR Infiniband |
| Performance | |
| Linpack Performance (Rmax) | 148,600 TFlop/s |
| Theoretical Peak (Rpeak) | 200,795 TFlop/s |
| Nmax | 16,473,600 |
| HPCG [TFlop/s] | 2,925.75 |
| Power Consumption | |
| Power: | 10,096.00 kW (Submitted) |
| Power Measurement Level: | 3 |
| Measured Cores: | 2,397,824 |
| Software | |
| Operating System: | RHEL 7.4 |
| Compiler: | XLC, nvcc |
| Math Library: | ESSL, CUBLAS 9.2 |
| MPI: | Spectrum MPI |



ECM 245

Arquitetura e
Organização de
Computadores

5ª-feira

07h40-09h20

Sala: H204

AULA 16

Vantagens

Expansibilidade

Baixo custo

Alta disponibilidade

Balanceamento de carga



ECM 245

Arquitetura e
Organização de
Computadores

5ª-feira

07h40-09h20

Sala: H204

AULA 16

Desvantagens

Manutenção de equipamento

Monitoração dos nós

Gargalos de troca de informações