

# **ARQUITETURA DE SISTEMAS DISTRIBUÍDOS**

**EVOLUÇÃO DA COMPUTAÇÃO E  
INTRODUÇÃO AOS SISTEMAS DISTRIBUÍDOS**

# Olá!

O profissional da área de tecnologia da informação atua diretamente no planejamento, na implementação e na implantação de soluções de TI nas organizações. Dessa forma, é importante que compreenda os conceitos, os benefícios, as características e as restrições de cada arquitetura. Nesta aula abordaremos esses tópicos, com alguns exemplos.

Ao final desta aula, o aluno será capaz de:

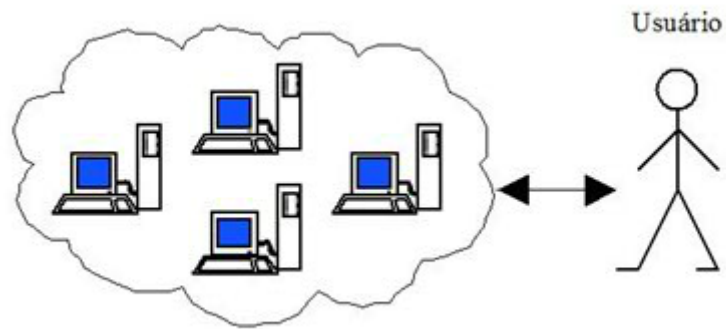
- 1- Entender as características da computação centralizada e da computação distribuída;
- 2- Conhecer as vantagens e desvantagens de cada arquitetura e a motivação para esta evolução;
- 3- Identificar o conceito de sistemas distribuídos.

## 1 Histórico

Primeiros computadores: grandes e caros.



Anos 50-60: spooling, multiprogramação.



Início dos anos 60: sistemas time sharing.



Final dos anos 60 e início dos anos 70: surgimento de redes de computadores.



A partir dos anos 70: inicia-se a pesquisa em sistemas distribuídos.



Desde seu início, a indústria de computadores tem se voltado para uma pesquisa interminável, em busca de ampliar, cada vez mais, seu poder no âmbito computacional. O Eniac podia executar 300 operações por segundo, sendo, tranquilamente, mil vezes mais rápido do que qualquer calculadora anterior a ele, e, mesmo assim, as pessoas ainda não estavam satisfeitas. Hoje temos máquinas um milhão de vezes mais rápidas do que o Eniac e, contudo, existe demanda para um poder computacional maior.



Não importa quanto poder computacional exista: ele nunca será suficiente.

No passado, a solução era fazer com que o relógio do processador fosse executado mais rápido (overclocking). Infelizmente, estamos começando a atingir alguns limites fundamentais na frequência do clock. De acordo com a teoria da relatividade de Einstein, nenhum sinal elétrico pode propagar mais rápido do que a velocidade da luz. Construir computadores com tamanho reduzido pode ser possível, mas, logo surge outro problema fundamental: a dissipação de calor. Quanto maior a frequência do relógio do processador, maior será a produção de calor produzida por ele; e quanto menor o computador, maiores serão os problemas associados à dissipação desse calor.

Uma maneira de aumentar a velocidade é fazer uso de computadores altamente paralelos. Essas máquinas são construídas com muitas CPUs (unidade central de processamento), objetivando processar, coletivamente, com muito mais poder computacional do que com apenas uma única CPU.



Por outro lado, com a evolução da internet, podemos analisar que um sistema com centenas de computadores espalhados pelo mundo não difere de um sistema de centenas de computadores em uma única sala, embora a latência e outras características técnicas sejam diferentes.

Neste contexto, colocar uma quantidade representativa de computadores em uma sala é até fácil, desde que se tenham recursos financeiros, espaço físico e uma infraestrutura mínima. O espaço deixará de ser um problema, se esses computadores estiverem espalhados ao redor do mundo. A preocupação surge quando queremos que eles se comuniquem uns com os outros para trabalhar em conjunto na solução de um único problema.

Os sistemas com múltiplos processadores caracterizam-se por possuir mais de uma CPU, interligadas, trabalhando em conjunto. Múltiplos processadores podem ser utilizados, simultaneamente, por diversos processos diferentes, e, com eles, novos problemas de concorrência são introduzidos, porque vários processos podem querer acessar um dispositivo ao mesmo tempo.



Esses sistemas podem ser divididos em:

- **Fortemente acoplados** – quando os processadores compartilham uma mesma memória principal.
- **Fracamente acoplados** – os diversos processadores/estações presentes no sistema utilizam sua memória local individualmente.
- **Sistemas centralizados** – Multiprocessador de memória compartilhada é um sistema de computador no qual duas ou mais CPUs compartilham acesso total a uma memória principal comum. Multiprocessadores são populares e atrativos, porque oferecem um modelo de comunicação simples, e a sincronização é possível mediante o emprego de diversas técnicas bem definidas. Uma desvantagem é que os multiprocessadores de grande porte são difíceis de construir e, por isso, são caros.
- **Sistemas distribuídos** – Uma solução alternativa que tem sido empregada com sucesso para solucionar esse problema é a utilização de multicomputadores, que são CPUs que não compartilham memória principal. Cada CPU tem sua própria memória e é gerenciada por um sistema operacional individualmente. Esses sistemas também são conhecidos como cluster – COWS (cluster of workstations - aglomerados de estações de trabalho).

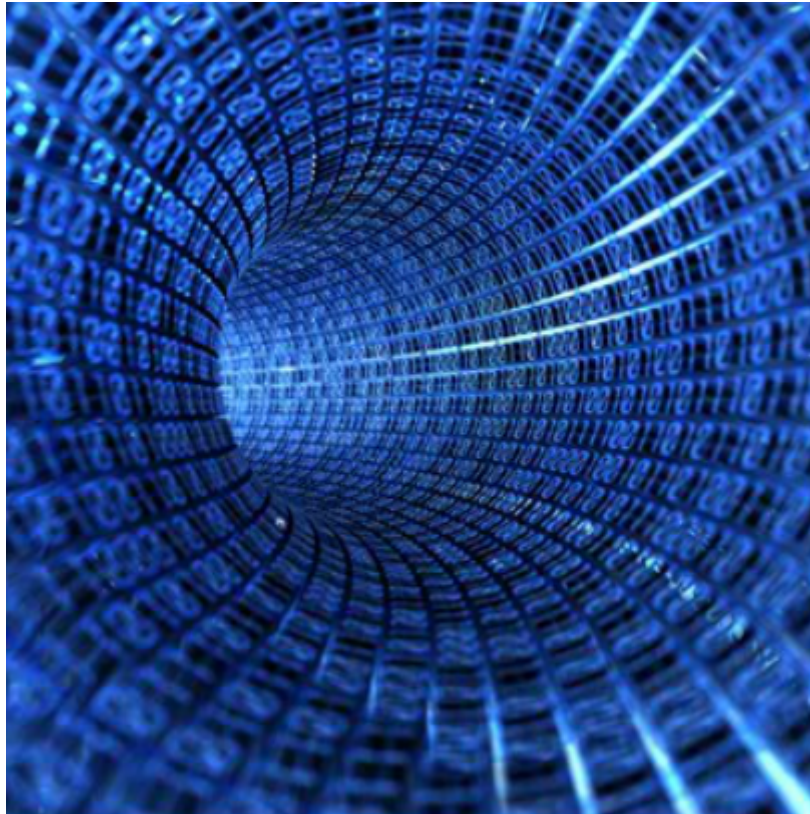
## 2 Por que computação distribuída?

Com a melhoria das tecnologias, o que se conseguia executar, algumas décadas atrás, somente com computadores que custavam milhões de dólares, hoje se consegue executar com computadores de baixo custo.



O segundo fator é o surgimento de redes de computadores de alta velocidade, em que informações podem ser transferidas entre computadores na faixa de microssegundos.

Como resultado, é possível conectar diversos computadores por meio de uma rede de alta velocidade para executar um sistema de computação colaborativo. Estes sistemas são geralmente chamados de sistemas distribuídos (SD).



### **Sistema distribuído permite uma nova forma de fazer ciência**

Teoria -> experimentos

Teoria -> simulações

### **Computação usada para modelar sistemas físicos**

#### **Vantagens**

- Possibilidade de repetição de eventos
- Manipulação de parâmetros
- Estudo de sistemas onde não há teorias exatas

#### **Problemas**

- **Muito grandes:** modelagem da terra/clima, simulações de reservatórios de petróleo, problemas com grandes escalas (cosmologia).
- **Muito pequenos:** projeto de remédios, projetos de chips, biologia estrutural, física de partículas.
- **Muito complexos:** física de partículas, dinâmica de fluidos, modelagem de comportamento de pessoas.

- **Muito caros:** produção e exploração de petróleo, simulação de acidentes.
- **Muito perigosos:** tolerância a falhas em aviões, teste de dispositivos nucleares, simulação de estratégias de defesa.

### Problemas que requerem alto desempenho computacional

- Modelagem climática
- Mapeamento do genoma
- Modelagem de semicondutores
- Dispersão de poluição
- Projetos farmacêuticos
- Modelagem de reatores nucleares
- Renderização de imagem

### Exemplo - previsão climática

Uma situação simplificada com seis variáveis: umidade, pressão, temperatura e velocidade do vento.

Memória: Terra discretizada em células de 1 km/1 km Na superfície da terra:  $5 \times 10^9$  células 24 bytes por célula  
 $= 10^{11} \approx 1$  TBytes

Processamento:

100 operações por célula

Intervalo 1 minuto

$$\frac{10^9 \text{ células}}{60s} \cdot 100 \frac{\text{operações}}{\text{célula}} \cong 8Gflops$$

### Aumento de desempenho

#### Limites físicos:

- Velocidade da luz.
- Miniaturização dos componentes.
- Isolamento e dissipação de calor.

Até meados de 1965 não havia nenhuma previsão real sobre o futuro do hardware, quando o então presidente da Intel, Gordon E. Moore, fez sua profecia, na qual o número de transistores dos chips teria um aumento de 100%, pelo mesmo custo, a cada período de 18 meses. Essa profecia tornou-se realidade e acabou ganhando o nome de Lei de Moore.

#### Solução: Paralelismo

- Execução simultânea de operações.
- Solução com melhor custo/benefício.



## 3 Evolução do processamento

### 3.1 Anos 70

Primeiras máquinas paralelas Illiac IV (64 processadores) foi um supercomputador construído pela Universidade de Illinois e financiado pelo governo dos EUA. Sua construção custou US\$ 31 milhões.



Figura 1 - Illiac IV

Fonte: [https://pt.wikipedia.org/wiki/Ficheiro:ILLIAC\\_4\\_parallel\\_computer.jpg](https://pt.wikipedia.org/wiki/Ficheiro:ILLIAC_4_parallel_computer.jpg)

### 3.2 Anos 80

- computadores vetoriais (Cray);
- máquinas paralelas comerciais para aplicações científicas (meteorologia, petróleo...);
- alto custo de desenvolvimento;
- pequena escala, dificuldade de programação.

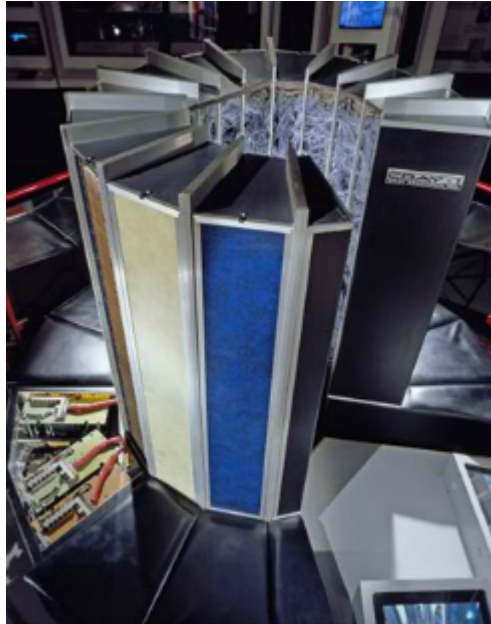


Figura 2 - Cray 1 - A Cray Research, empresa de Seymour Cray, o primeiro supercomputador vetorial (Cray), com altíssima velocidade de processamento e grande capacidade de memória, empregado em pesquisas científicas e militares.

Fonte: <http://museudoscomputadores.blogspot.com/>

### 3.3 Anos 90 aos dias de hoje

- multiprocessadores escaláveis;
- redes de estações de trabalho;
- computação distribuída;
- aplicações cliente/servidor;
- objetos distribuídos;
- clusters;
- grids.

Com o surgimento de malhas computacionais (grids), em que as aplicações executam em ambientes distribuídos, compostos por um grande número de máquinas heterogêneas, administradas por diferentes instituições e conectadas à Internet.



Figura 3 - Computação em grade (grid computing)

Fonte: <http://www.adarshpatil.com/newsite/grid.htm>

## 4 Paralelismo X Computação paralela

### Paralelismo

- Projeto de uma CPU
- Projeto de uma arquitetura paralela
- E/S sobreposta ao processamento

### Computação paralela

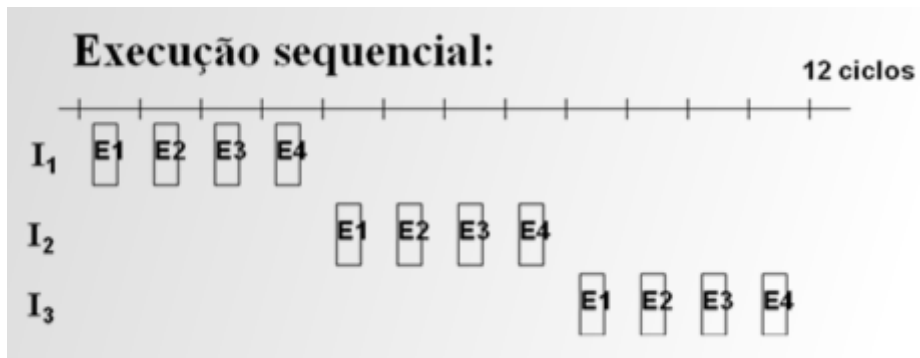
- Coleção de elementos de processamento
- Trabalhando em conjunto para a realização de uma tarefa

### Paralelismo

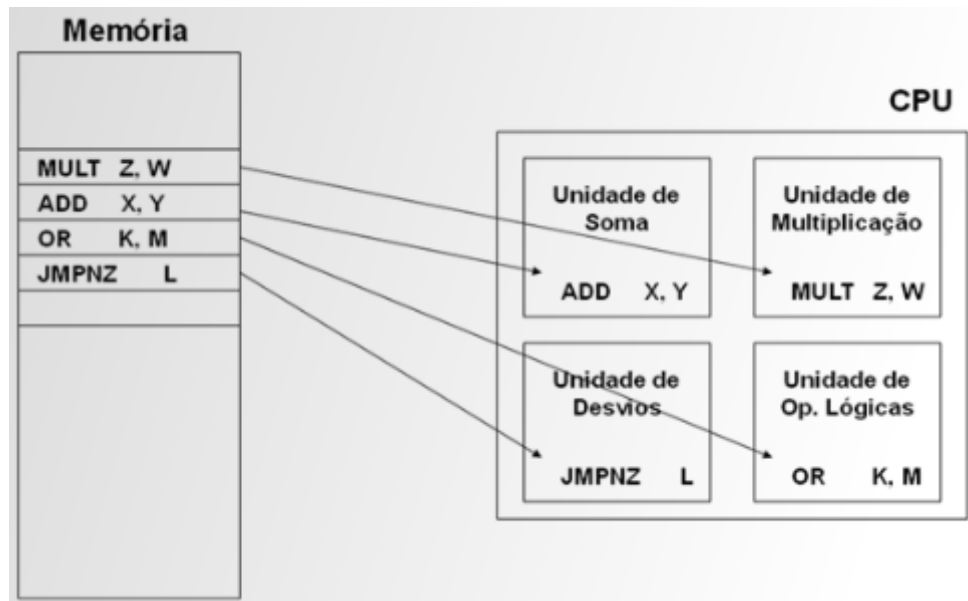
Dentro de um processador:



## 5 Arquiteturas sequenciais



## 6 Arquiteturas superescalares



## 7 Arquiteturas SMT

SMT = Simultaneous Multithreading

Múltiplos threads despacham múltiplas instruções no mesmo ciclo

Aumenta a possibilidade de manter todas as unidades funcionais ocupadas

## 8 Arquiteturas pipeline

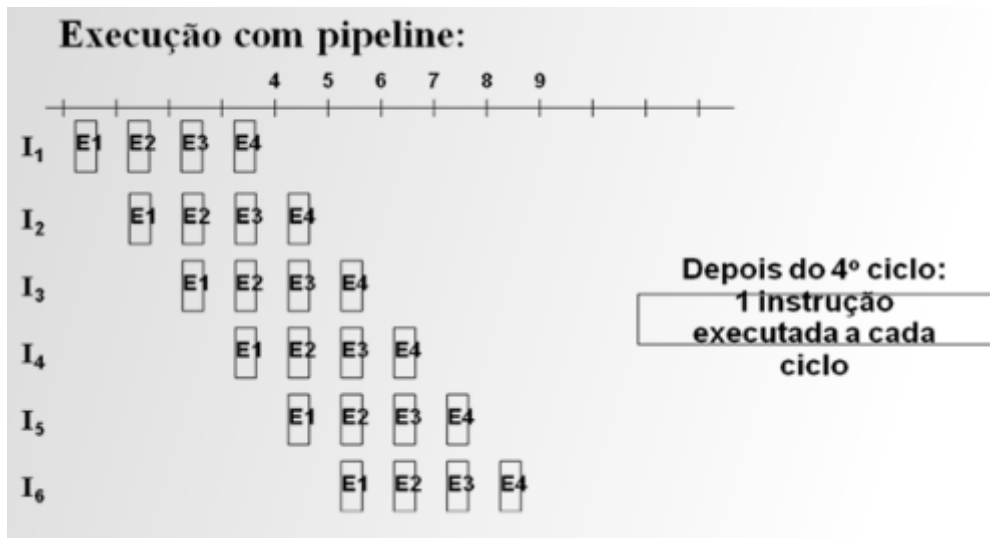


Figura 4 - Pipeline

## 9 Arquiteturas VLIW

VLIW = Long Instruction Word

Também requer unidades funcionais operando em paralelo

Também permite execução simultânea de instruções

VLIW diferente de superescalar: VLIW - as dependências são resolvidas em tempo de compilação

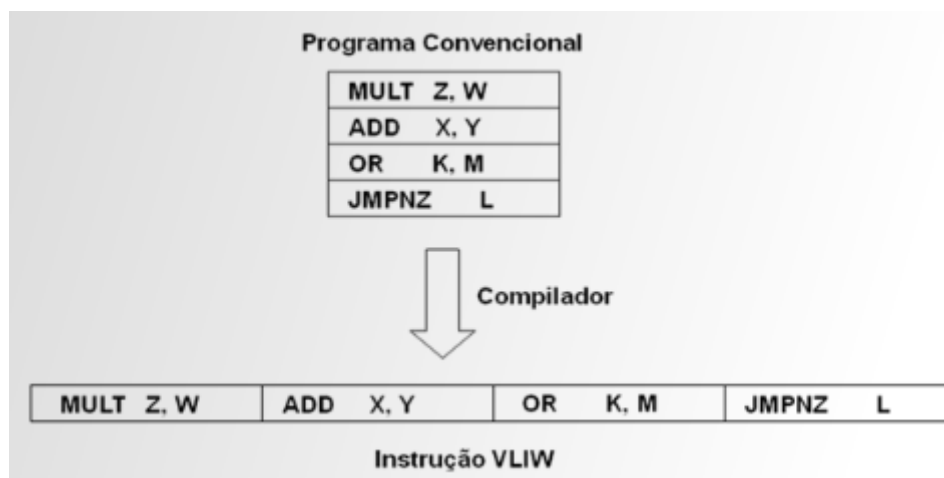


Figura 5 - Arquitetura VLIW



## 10 Clusters

**Definição:** Conjunto de computadores independentes interconectados, usados como recurso unificado para computação.

Sistema com imagem única (SSI Single System Image)

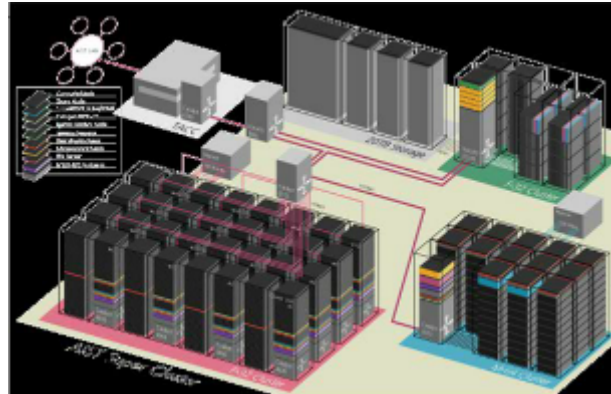


Figura 6 - Computação em cluster

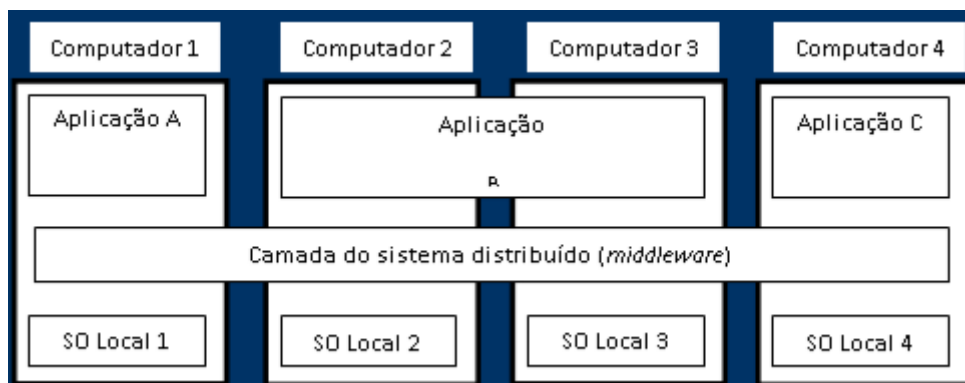


Figura 7 - Sistema distribuído organizado como middleware (SSI)

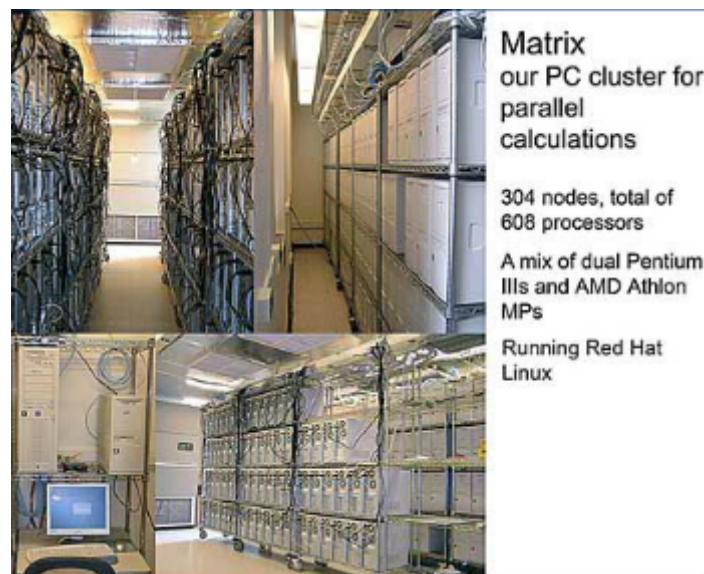
## 10.1 Por que clusters?

- **Custo/benefício** - redução de custo para se obter processamento de alto desempenho, utilizando máquinas de baixo custo.
- **Escalabilidade** - possibilidade de inclusão de novos componentes, que sejam adicionados à medida que cresça a carga de trabalho.
- **Alto desempenho** - possibilidade de resolver problemas complexos através de programação e processamento paralelo, reduzindo o tempo para a solução do problema.
- **Independência de fornecedores** - utilização de hardware aberto, software de uso livre e independência de fabricantes e licenças de uso.
- **Tolerância a falhas** - o aumento de confiabilidade do sistema como um todo, caso alguma parte falhe.

## 10.2 Cluster Beowulf

Em 1994, o Projeto Earth and Space Sciences (ES) da NASA construiu o primeiro cluster Beowulf para fornecer computação paralela com o intuito de resolver problemas envolvidos em aplicações da ESS.

Uma característica desta arquitetura é que é constituída por diversos nós escravos e gerenciada por um só computador.



O primeiro cluster de PC's (Beowulf) foi construído em 1994 por dois pesquisadores do Goddard Space Flight Center (NASA) Thomas Sterling e Don Becker. Naquele momento esses pesquisadores necessitavam de uma solução de baixo custo e que processasse na ordem de 1 Gflop, o que significa um bilhão de operações em ponto flutuante por segundo, já que preço de um supercomputador com este nível de desempenho era de aproximadamente um milhão de dólares. A idéia original era utilizar processadores mais baratos, para tanto, este cluster foi configurado inicialmente com 16 máquinas utilizando processadores 486DX4, placas de rede

Ethernet e sistema operacional Linux. Este primeiro cluster atingiu a marca de 70 Mflops, com um custo estimado em 10% do valor de mercado para uma máquina de desempenho similar. Este projeto obteve tanto sucesso que o termo Beowulf é utilizado até hoje como uma classe de clusters.



Estas máquinas foram montadas e programadas de forma que possibilitassem a paralelização/divisão das tarefas, buscando atingir um poder de processamento equivalente a um supercomputador da época, por uma fração do preço. Este primeiro cluster chegou 70 Mflops, com um custo estimado em 1 /10 do valor cobrado pelo mercado para um sistema de desempenho similar. Tal projeto fez tanto sucesso que o termo Beowulf foi estendido a uma classe de clusters que são utilizados ainda hoje.

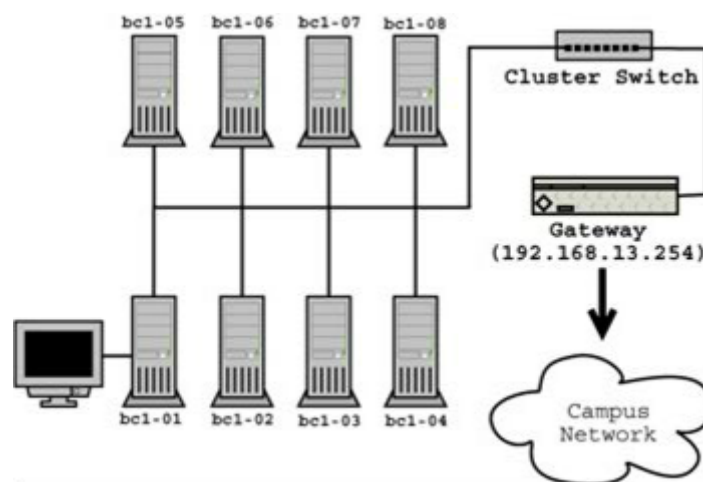


Figura 8 - Diagrama esquemático de um cluster Beowulf

Fonte: <http://etutorials.org/Linux+systems/cluster+computing+with+linux/Part+I+Enabling+Technologies/Chapter+5+Configuring+and+Tuning+Cluster+Networks/5.4+Simple+Cluster+Configuration+Walkthrough/>

## 11 Computação centralizada

**Mainframe:** termo utilizado para se referenciar a um grande computador, normalmente produzido por uma grande empresa. O nome tem origem na forma com que estes computadores eram construídos. Todos os componentes (processador, memória...) do computador principal (main) são colocados em uma única estrutura (frame).

### Características

- Sistemas multiusuários
- Sistemas proprietários -> hardware, software, rede
- Instalação e manutenção feita pelo fabricante -> confiabilidade X custo

## 12 Microcomputadores e redes de computadores

Ampliação do parque computacional, em função de:

- processadores mais rápidos e mais baratos;
- redes mais rápidas e acessíveis;
- liberdade de escolha;
- menor custo de manutenção;
- necessidade inerente de conectividade;
- aplicação básica: compartilhamento de recursos;

**Evolução:** Os terminais foram sendo substituídos pelos primeiros microcomputadores que começavam a ficar obsoletos. Em geral, o uso de um programa emulador de terminais e de uma unidade de disquete era suficiente para que um simples PC-XT executasse essa tarefa, uma vez que só precisaria executar o emulador. A partir deste ponto, o micro passaria a se comportar como um terminal. Em alguns casos, era necessário o uso de uma placa que compatibilizasse a forma de comunicação serial entre os dois computadores.

## 13 Sistemas distribuídos

Utilização das redes de computadores (locais e de longa distância) para execução colaborativa e cooperativa de aplicações e não somente para compartilhamento de recursos.

**Sistema distribuído = computadores + rede + aplicação**

**Conceito:** É um sistema em que os computadores estão conectados em rede e coordenam suas ações através de troca de mensagens.

### Definições de sistemas distribuídos

- **Colouris:** "Um sistema no qual os componentes de hardware ou software, localizados em computadores interligados em rede, se comunicam e coordenam suas ações apenas enviando mensagens entre si."

- **Tanenbaum:** "Um sistema distribuído é um conjunto de computadores independentes que se apresenta a seus usuários como um sistema único e coerente."
- **Silberschatz:** "Coleção de processadores que não compartilham memória ou relógio."

#### Principais motivações

- necessidade pelo compartilhamento de recursos;
- recurso pode ser um serviço, arquivo, banco de dados, streaming de vídeo etc.

#### Características de um sistema distribuído

- baixo acoplamento e atrasos na comunicação;
- processos em sistemas computacionais distintos com probabilidade de falhas;
- comunicação geralmente não confiável, onde existem atrasos, variação de atrasos, perdas e, em alguns casos, baixas larguras de banda;
- dificuldade em definir a ordem dos eventos e estado global do sistema, uma vez que a comunicação acontece pela troca de mensagens;
- ambiente geralmente marcado pela heterogeneidade.

## 14 Comparação com sistemas centralizados

#### Vantagens

- Melhor relação preço/desempenho
- Capacidade de crescimento incremental (escalabilidade)
- Tolerância a falhas

#### Desvantagens

- Falta de padronização para desenvolvimento de software
- Falta de uma divisão clara entre sistema/aplicação
- Latência e possibilidade de congestionamento na rede
- Redução da segurança

## 15 Desafios da computação distribuída

- Ausência de fonte comum de tempo (relógio global)
- Ausência de memória compartilhada
- Compartilhamento de recursos

## O que vem na próxima aula

Na próxima aula, você vai estudar:

- Identificar o uso de aplicações distribuídas;
- Conhecer as principais aplicações atuais;
- Compreender como a computação distribuída pode contribuir para TI verde.

# CONCLUSÃO

Nesta aula, você:

- Entendeu a necessidade da utilização da computação de alto desempenho nos ambientes corporativos;
- Conheceu as características da computação centralizada e da computação distribuída;
- Identificou as vantagens e desvantagens de cada arquitetura e a motivação para esta evolução.