

# Sistemas Distribuídos

Histórico, terminologia, sistemas centralizados, distribuídos e paralelos

1

Universidade Estácio de Sá  
Professor Welsing M. Pereira



## Breve Histórico



- **1945 a 1985:** presença de computadores grandes e de custo elevado que trabalhavam de maneira independente por não haver uma forma confiável de conectá-los.
- **1985 em diante:** o avanço tecnológico provocou mudanças importantes em 2 pontos principais:
  - 1º Desenvolvimento de microcomputadores com maior poder computacional (CPUs de 8, 16, 32 e 64 bits) a um preço mais atraente.
  - 2º Surgimento das redes de alta velocidade e, conseqüentemente, dos *SISTEMAS DISTRIBUÍDOS*.
    - a. LANs (Local Area Networks): permitindo a conexão de centenas de computadores e a troca de pequenas quantidades de informação em alguns microssegundos .
    - b. WANs (Wide Area Networks) permitindo a conexão de milhões de computadores pelo mundo com velocidade de transmissão dos dados entre 64kbts/seg até Gigabits/seg.

2

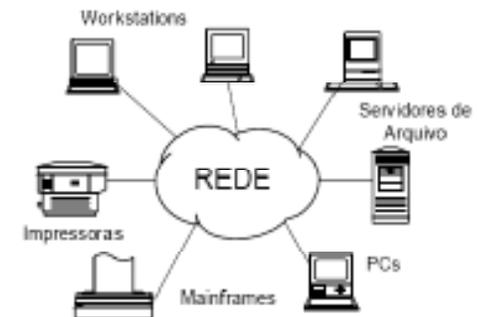
## Fundamentos



- O que é um sistema distribuído?
- Vejamos as definições de Coulouris e Tanenbaum:
  - Coleção de computadores autônomos interligados através de uma rede de computadores e equipados com software que permita o compartilhamento dos recursos do sistema: hardware, software e dados [Coulouris];
  - Coleção de computadores independentes que se apresenta ao usuário como um sistema único e consistente [Tanenbaum].
- Portanto, podemos definir um SD como uma coleção de computadores independentes que aparenta aos usuários do sistema como se fosse um único computador. Nesta definição devemos considerar dois aspectos:
  - Hardware: as máquinas são autônomas.
  - Software: os usuários vêem o sistema como uma única máquina.
- No SD os computadores não compartilham memória nem relógio. Em vez disso, cada processador tem sua própria memória local e os processadores se comunicam por meio de redes de interconexão.

3

## Fundamentos



- Na ilustração apresentada acima, além dos componentes listados, é necessário em um SD que exista um mecanismo de controle para gerenciar as atividades e administrar, de forma eficiente, os recursos disponíveis no ambiente.

4

## Características de um SD



- A seguir destacamos algumas características dos sistemas distribuídos. Nem todos os SDs apresentam as mesmas propriedades; isso pode depender do seu propósito.
  - Compartilhamento de recursos: impressoras, discos, banco de dados, arquivos, etc.
  - Concorrência e paralelismo: o acesso aos recursos devem ser sincronizados.
  - Escalabilidade.
  - Tolerância a Falhas: falhas independentes.
  - Alta Disponibilidade
  - Não existe um clock global
  - Transparência em relação as falhas, a localização de recursos, acesso, replicação, escalabilidade, concorrência, mobilidade, desempenho, etc.

5

## Vantagens de um SD



- Podemos listar algumas vantagens dos SDs sobre os sistemas centralizados, dentre elas destacamos:
  - Compartilhamento de dados: base de dados comum;
  - Compartilhamento de dispositivos: acesso compartilhado a periféricos;
  - Comunicação: torna-se mais simples e mais rápida a comunicação entre pessoas. Além disso, é possível: transferência de arquivos entre nós, login remoto, etc;
  - Flexibilidade: dividir a carga de trabalho entre os nós da rede;
  - Confiabilidade: se um nó falha os demais poderão continuar operando;
  - Velocidade de computação: maior poder computacional obtido através de concorrência. Há a possibilidade de distribuir uma computação particionada a vários nós para executarem concorrentemente;
  - Performance a baixo custo: preço baixo dos PCs;
  - Escalabilidade: aumentar o número de nós na medida da necessidade.

6

## Desvantagens



- Podemos listar as desvantagens de um SD sobre um sistema centralizado em relação aos seguintes pontos:
- Software:
  - a. Falta de experiência no projeto, implementação e uso de software distribuído.
  - b. Quanto o usuário deve conhecer sobre a distribuição?
  - c. Qual o trabalho que deve ser feito pelo sistema e pelo usuário?
- Redes de comunicação:
  - a. Perda de mensagens na rede.
  - b. Sobrecarga na rede gerada pelas comunicações entre nós.
  - c. Problema no dimensionamento da rede.
- Segurança:
  - Apesar do compartilhamento de dados ser uma vantagem devemos cuidar da segurança das informações.

7

## Sistemas Centralizados



- Inicialmente, todos os sistemas de computadores eram centralizados, ou seja, compostos por um único processador, uma memória, alguns periféricos e terminais. Num sistema centralizado, todos os dispositivos de armazenamento secundário e demais periféricos estão ligados a uma única CPU. O sistema operacional, normalmente formado por um grande núcleo monolítico, permite que os usuários compartilhem os recursos disponíveis. O sistema de arquivos, como um subsistema do sistema operacional, também é centralizado, atendendo às requisições de todos os usuários. O sistema de arquivos e seus usuários são implementados como processos, executados numa mesma máquina.

8

## Sistemas Distribuídos, Paralelos ou de Alto Desempenho



- Entende-se por Computação de Alto Desempenho como a área da computação preocupada com a criação de condições para que processamentos de elevada carga computacional possam ser executados em bem menos tempo. Fica claro então que uma condição necessária para a obtenção do alto desempenho é a existência de vários elementos de processamento no sistema (sistemas com múltiplos processadores), permitindo que vários programas sejam executados ao mesmo tempo (paralelismo) ou que um mesmo programa seja subdividido em partes para serem executadas simultaneamente em mais de um processador (processamento paralelo).
- Um fator-chave no desenvolvimento de sistemas operacionais com múltiplos processadores é a forma de comunicação entre as CPU's e o grau de compartilhamento da memória e dos dispositivos de entrada e saída. Em função desses fatores, podemos classificar os sistemas em fortemente e fracamente acoplados.

9

## Sistemas Distribuídos, Paralelos ou de Alto Desempenho



- **Sistemas fortemente acoplados** são sistemas em que os vários elementos de processamento compartilham o barramento, o clock e, às vezes, a memória e os dispositivos periféricos, isto implica em máquinas de alto custo e grande velocidade de processamento, o oposto de **Sistemas fracamente acoplados**, quando os processadores e memórias estão ligados por um barramento diferente do local, como por exemplo, através de uma conexão de rede, e cada um dos sistemas funciona de forma independente, trazendo uma grande redução no custo do sistema.
- Geralmente, sistemas fortemente acoplados são usados como sistemas paralelos (trabalhando em um único problema), onde é necessário que um barramento seja muito rápido e consiga transferir uma grande quantidade de dados em um determinado tempo, enquanto que sistemas fracamente acoplados são utilizados como um sistema distribuído (trabalhando em muitos problemas não relacionados), onde a comunicação entre os processadores e memórias se dá através de troca de mensagens pela rede, embora isso não seja sempre verdade.

10

## Sistemas Distribuídos, Paralelos ou de Alto Desempenho



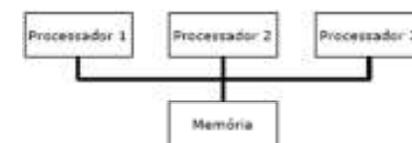
- A Computação Paralela e Sistemas Distribuídos são áreas raramente estudadas em separado, e às vezes seus conceitos se confundem nas mais diversas situações e aplicações, entretanto, são igualmente independentes. É possível existir um Sistema Distribuído que não utilize Paralelismo, como por exemplo, um grupo de computadores que alocam as CPUs de acordo com a demanda. Nesse caso, não há paralelismo, há apenas distribuição de carga para máquinas que estejam mais ociosas. É também possível existir um ambiente que implementa Computação Paralela, mas não é necessariamente um Sistema Distribuído, como em uma única máquina com 2 ou mais processadores. As tarefas desse sistema são executadas em paralelo de acordo com a disponibilidade dos processadores, entretanto, não se trata de um Sistema Distribuído, mas sim, de uma máquina multiprocessada. Para esclarecer melhor, estudaremos a seguir os conceitos de cada sistema separadamente.

11

## Sistemas Paralelos



- Os sistemas paralelos são caracterizados pela presença de mais de um processador compartilhando o barramento, o clock e, às vezes, a memória e os dispositivos periféricos. Tais sistemas são chamados de **fortemente acoplados**.



Sistema Paralelo com memória compartilhada

O principal objetivo desses sistemas é oferecer maior capacidade de processamento de modo que o tempo de execução das aplicações possa ser diminuído. Quando aumentamos o número de processadores de 1 para  $n$  esperamos que o tempo de execução de uma aplicação seja dividido por  $n$ , mas isso é difícil acontecer. Este fato pode ser explicado, pois há um esforço necessário para manter todas as partes trabalhando corretamente. Esse esforço, mais a disputa por recursos compartilhados, diminui o ganho esperado dos processadores adicionais.

12

## Sistemas Paralelos



- Uma das vantagens desses sistemas é economizar dinheiro com periféricos, memórias, fontes de alimentação, etc. Outra vantagem é o aumento da confiabilidade, pois, a falha de um processador não vai interromper o sistema, apenas reduzirá a velocidade (tolerância a falhas).
- Um esquema muito utilizado é o de multiprocessamento simétrico, onde todos os processadores são iguais, no qual cada processador executa uma cópia idêntica do sistema operacional de forma concorrente. Alguns sistemas utilizam multiprocessamento assimétrico, no qual a cada processador é atribuída uma tarefa específica. Um processador mestre controla o sistema; os outros processadores procuram o mestre para receber instruções ou tem tarefas predefinidas. O processador mestre escalona e aloca trabalho para os processadores escravos.
- Apesar da computação paralela ter como objetivo principal aumentar o desempenho na implementação de problemas específicos, vários trabalhos tem sido desenvolvidos com o objetivo de utilizar os sistemas distribuídos para computação paralela. A idéia básica é ter um grupo de computadores interligados, funcionando como os elementos de processamento de uma máquina paralela.

13

## Sistemas Distribuídos



- Um sistema distribuído é uma coleção de computadores que não compartilham memória nem relógio. Em vez disso, cada processador tem sua própria memória local. Os processadores se comunicam entre si por meio de várias redes de comunicação, ou seja, um sistema distribuído é uma coleção de processadores **fracamente acoplados**, interconectados por uma rede de comunicação em um **software fortemente acoplado** com o objetivo de criar a ilusão de que toda a rede de computadores nada mais é do que um único sistema de tempo compartilhado, em vez de um conjunto de máquinas distintas.



Sistema Distribuído (memória não-compartilhada)

14

## Sistemas Distribuídos



- A ausência de memória compartilhada exige que a interação entre processadores ocorra de uma forma distinta do ambiente centralizado: ao invés de variáveis ou arquivos compartilhados utiliza-se troca de mensagens.
- Um sistema distribuído reúne sistemas fisicamente separados e possivelmente heterogêneos em um sistema único coerente fornecendo aos usuários acesso aos vários recursos mantidos pelo sistema permitindo compartilhamento de recursos, melhor confiabilidade e maior velocidade de computação, se uma determinada computação puder ser particionada em uma série de subcomputações que executam concorrentemente ou também se, um determinado site estiver sobrecarregado de jobs, alguns deles podem ser movidos para outros sites menos sobrecarregados.
- O motivo principal para a criação e desenvolvimento de sistemas distribuídos foi, inicialmente, a necessidade de se compartilhar recursos, normalmente de alto custo e separados fisicamente.

15

## Diferenças entre Sistemas Distribuídos e Sistemas Paralelos



Sistemas Distribuídos	Sistemas Paralelos
Arquitetura baseada em uso de redes	Os processadores estão todos localizados numa mesma placa mãe, ou em poucas
Fracamente acoplados	Fortemente acoplados
São mais imprevisíveis devido ao uso da rede e a falhas	Comportamento é mais previsível
São bastante influenciados pelo tempo de comunicação pela rede	O tempo de troca de mensagens pode ser desconsiderado
Não possuem limitações em número de máquinas	Possuem limitações em processadores
Cada processador com sua memória	Processadores geralmente compartilham memória
Tolerante a falha, pois caso algum computador seja retirado da rede, o funcionamento não será prejudicado	Não é tolerante a falha, no caso de um processo ficar retido no processador, e existirem processos que dependam deste
Troca de dados somente por mensagens	Troca de dados pode ser realizada utilizando Memória
Aumento do processamento é mais fácil, pois basta adicionar mais uma máquina na rede	Aumento do processamento depende dos avanços da tecnologia e hardware
Alta disponibilidade e compartilhamento de informações	Alto desempenho

16