

Computação Distribuída

Fundamentos de Sistemas
Distribuídos



O que são sistemas distribuídos?

- São sistemas em que pelo menos um dos componentes está localizado em uma **rede de computadores** e coordenam suas ações através da **troca de mensagens**
- O **compartilhamento de recursos** é a motivação principal para construir sistemas distribuídos

Os desafios dos Sistemas Distribuídos

- Heterogeneidade dos componentes
- Facilidade em trocar ou adicionar novos componentes
- Segurança
- Escalabilidade
- Gerenciamento de falhas
- Concorrência
- Transparência

Os desafios dos Sistemas Distribuídos

- **Heterogeneidade:**
 - **redes:** protocolos, velocidade, confiabilidade;
 - **hardware de computador:** representação de dados, arquiteturas;
 - **sistemas operacionais:** APIs, programação em multitarefa;
 - **linguagens de programação:** tipos e estruturas de dados, gerenciamento de memória, passagem de parâmetros;
 - **implementações de diferentes desenvolvedores:** falta de aderência a padrões.

Características dos Sistemas Distribuídos

- Os computadores podem estar em continentes diferentes, no mesmo prédio ou na mesma sala, levando a determinadas consequências:
 - **Concorrência:** é a norma em redes de computadores.
 - Um usuário executa determinados processos em um computador e outro usuário faz coisas diferentes em outro computador. Porém, ambos podem compartilhar recursos comuns
 - Um dos principais tópicos em SD
 - **Ausência de um relógio global:** os programas cooperam entre si através da troca de mensagens.
 - Como a comunicação depende de fatores externos à aplicação, como atrasos e desconexões, problemas com relação ao tempo de execução podem ocorrer

Características dos Sistemas Distribuídos

- O Compartilhamento de Recursos Distribuídos envolve:
 - Componentes de hardware:
 - Discos, impressoras...
 - Componentes de software:
 - Arquivos, bancos de dados...

Características dos Sistemas Distribuídos

- SDs devem possuir Extensibilidade
 - Extensões de hardware
 - Periféricos, memória, interfaces de comunicação
 - Extensões de software
 - Funções de SO, protocolos de comunicação
- Principais interfaces são públicas (system calls)
- Existe um mecanismo uniforme de comunicação entre processos

Características dos Sistemas Distribuídos

- Existe Concorrência em SD
 - Mais de um processo em execução a cada instante:
 - Atividades separadas de usuários
 - Independência de recursos
 - Localização de processos servidores em computadores distintos
 - Acesso concorrente a recursos compartilhados requer sincronização

Características dos Sistemas Distribuídos

- Tolerância a falhas
 - Falhas de hardware e software:
 - Programas param ou produzem resultados errados
 - Abordagens:
 - Redundância de hardware
 - Ex : banco de dados replicado em diversos servidores
 - Recuperação por software:
 - Manter dados permanentes sempre consistentes

Características dos Sistemas Distribuídos

- Propriedades Críticas de SD:

- Transparência
- Flexibilidade
- Confiabilidade
- Desempenho
- Escalabilidade

Propriedade Crítica

- **Transparência**

- É esconder do usuário e do programador de aplicações a separação de componentes em um sistema distribuído, aparentando ser um sistema centralizado
- Formas de transparência
 - **Acesso:** operações de acesso a objetos de informação são idênticas para objetos locais e remotos
 - **Localização:** acesso a um objeto ocorre sem que seja necessário o conhecimento de sua localização
 - **Concorrência:** processos operam concorrentemente usando objetos de informação comuns sem interferência entre eles

Propriedade Crítica

- **Transparência**

- É esconder do usuário e do programador de aplicações a separação de componentes em um sistema distribuído, aparentando ser um sistema centralizado
- Formas de transparência
 - **Replicação:** várias instâncias de um objeto de informação são usadas sem requerer o conhecimento das réplicas pelos usuários e aplicações
 - **Falha:** mascaramento de falhas de hardware e software
 - **Migração:** movimento de objetos de informação dentro do sistema não afeta a operação de usuários e aplicações
 - **Desempenho:** reconfiguração do sistema para melhorar desempenho conforme a carga varia
 - **Escala:** o sistema e as aplicações podem expandir em escala sem requerer modificações na estrutura do sistema ou nos algoritmos das aplicações

Propriedade Crítica

- **Transparência**

- A transparência depende também da distância entre os servidores e os clientes do SD
- Existem atrasos na rede:
 - Congestionamento
 - Atrasos relativos ao roteamento
 - Replicação entre servidores pode causar atrasos na rede
 - Novas rotas devido a um servidor inoperante podem causar mais atrasos ao cliente

Propriedade Crítica

- **Flexibilidade**

- Modelos de estrutura de SD:

- **Núcleo monolítico**

- inclui gerenciamento de arquivos, diretórios e processos

- **Micro núcleo**

- mecanismo para comunicação entre processos
 - gerenciamento básico de memória
 - gerenciamento de processos a baixo nível
 - operações de entrada/saída a baixo nível

Propriedade Crítica

- **Flexibilidade**
 - **Modularidade => Flexibilidade**
 - Serviços estão disponíveis a todos os clientes, independentemente de localização
 - Serviços podem ser reparados sem causar parada total do sistema
 - Os próprios usuários podem adicionar novos serviços

Propriedade Crítica

- **Confiabilidade**

- Aspectos:
 - Disponibilidade: fração do tempo em que o sistema pode ser usado
 - Exatidão: replicação *vs* consistência
 - Segurança: Como um servidor pode verificar a origem de uma mensagem?
 - Tolerância a falhas: replicação *vs* desempenho

Propriedade Crítica

- **Escalabilidade**

- Quantidade de trabalho envolvido no processamento de qualquer requisição de acesso a um recurso compartilhado independe do tamanho da rede
- Técnicas:
 - Replicação
 - Caching
 - Servidores múltiplos

Propriedade Crítica

- **Escalabilidade**

- Evitar:
 - Componentes centralizados
 - ex : um único servidor de correio eletrônico para todos os usuários
 - Tabelas centralizadas
 - ex : uma única lista telefônica on line
- Algoritmos centralizados
 - ex : roteamento baseado em informação completa

Algoritmos Distribuídos

- Nenhuma máquina conhece o estado global
- Decisões são baseadas somente em informação local
- A parada de uma máquina não arruína o algoritmo
- Não há qualquer suposição quanto a existência de tempo (relógio) global

Elementos de Sistemas Distribuídos

- Sistema de nomes
- Comunicação
- Estrutura de software
- Alocação de carga de trabalho
- Manutenção de consistência

Elementos de Sistemas Distribuídos

- **Sistema de nomes**

- Nomes permitem que recursos sejam compartilhados
- Nomes de recursos devem ser independentes de sua localização
- O esquema de nomes deve escalar bem
- Um sistema de interpretação de nomes deve ser acessível por programa

DNS

Domain Name System

1. Usuário abre um browser e digita um nome de domínio
2. Seu computador pergunta ao DNS local pelo nome de domínio
3. O DNS local pergunta aos servidores raiz quem é responsável pelo domínio
4. O servidor raiz indica o servidor DNS responsável pelo domínio
5. O DNS local pergunta ao DNS responsável qual o endereço do nome de domínio
6. O DNS responsável traduz o nome de domínio p/ IP respondendo ao DNS local
7. O DNS local responde ao seu computador o endereço IP
8. Seu computador conecta ao servidor desejado

Elementos de Sistemas Distribuídos

- **Comunicação**

- O sucesso de um SD depende muito do desempenho/confiabilidade das técnicas de comunicação usadas em sua implementação
- Dilema: otimizar implementação da comunicação e prover alto nível do modelo de programação dessa comunicação
- As redes de comunicação são responsáveis pela efetividade de um sistema distribuído (ou seu fracasso)

Elementos de Sistemas Distribuídos

- **Estrutura de Software**

- Extensibilidade requer componentes de software com interfaces bem definidas
- Um serviço é um gerenciador de objetos de um certo tipo e sua interface é um conjunto de operações
- Novos serviços devem interoperar com serviços existentes e não duplicar suas funções

Elementos de Sistemas Distribuídos

- **Alocação de carga de trabalho**

- Otimização do uso de:
 - capacidade de processamento
 - capacidade de comunicação
 - recursos da rede em geral
- Objetivo: bom desempenho

Elementos de Sistemas Distribuídos

- **Manutenção de Consistência**

- **Consistência de atualização:** atomicidade como meio de atualização instantânea de muitos elementos
- **Consistência de replicação:** cópias de um mesmo recurso devem ser “idênticas”
- **Consistência de cache:** modificações em um cliente devem ser propagadas ao gerenciador e aos demais clientes
- **Consistência de falha:** deve se evitar falhas múltiplas (em cascata); isolamento de falhas
- **Consistência de relógio:** relógios físicos (sincronização aproximada) e relógios lógicos (timestamps em mensagens)
- **Consistência de interface de usuário:** atrasos devido a comunicação podem causar visão inconsistente de aplicações gráficas

Segurança em Sistemas Distribuídos

- SD's possuem normalmente milhares de pessoas conectadas aos seus serviços.
- Falhas de segurança podem acontecer
 - Usuários desatentos ou mal intencionados

Segurança em Sistemas Distribuídos

- Softwares mais antigos não foram concebidos para serem processados fora do ambiente interno das empresas e, não raro, quando usados de fora da empresa apresentam toda sorte de vulnerabilidades
- Quando sistemas legados são portados para SDs , podem ocorrer problemas de segurança como:
 - Política de segurança local geralmente não é aplicável a ambientes distribuídos
 - Senhas locais podem ter sido armazenadas sem criptografia

Segurança em Sistemas Distribuídos

- O rastreamento de comunicações pode traçar o perfil de preferências de um usuário em especial
 - Isso viola a privacidade, principalmente quando isso é feito sem o prévio conhecimento do usuário
 - Softwares de monitoramento ou arquivos de registro sobre preferências pessoais do usuário (como cookies) tanto podem ser positivas para aquele usuário que deseja obter informações de acordo com sua preferência, como podem ser invasivas
 - Não existe uma garantia de que os dados e preferências do usuário não vão ser expostos ou comercializados por usuários maliciosos

Sistemas Distribuídos Abertos

- Um sistema distribuído é aberto porque oferece serviços de acordo com regras padronizadas que descrevem a sintaxe e a semântica desses serviços
- Numa rede de computadores há regras padronizadas que governam o formato, o conteúdo e o significado das mensagens enviadas e recebidas
- Essas regras são formalizadas em protocolos, sendo especificadas por meio de uma interface IDL (Interface Definition Language ou Linguagem de Definição de Interface)
 - Uma IDL é normalmente utilizada em Middlewares

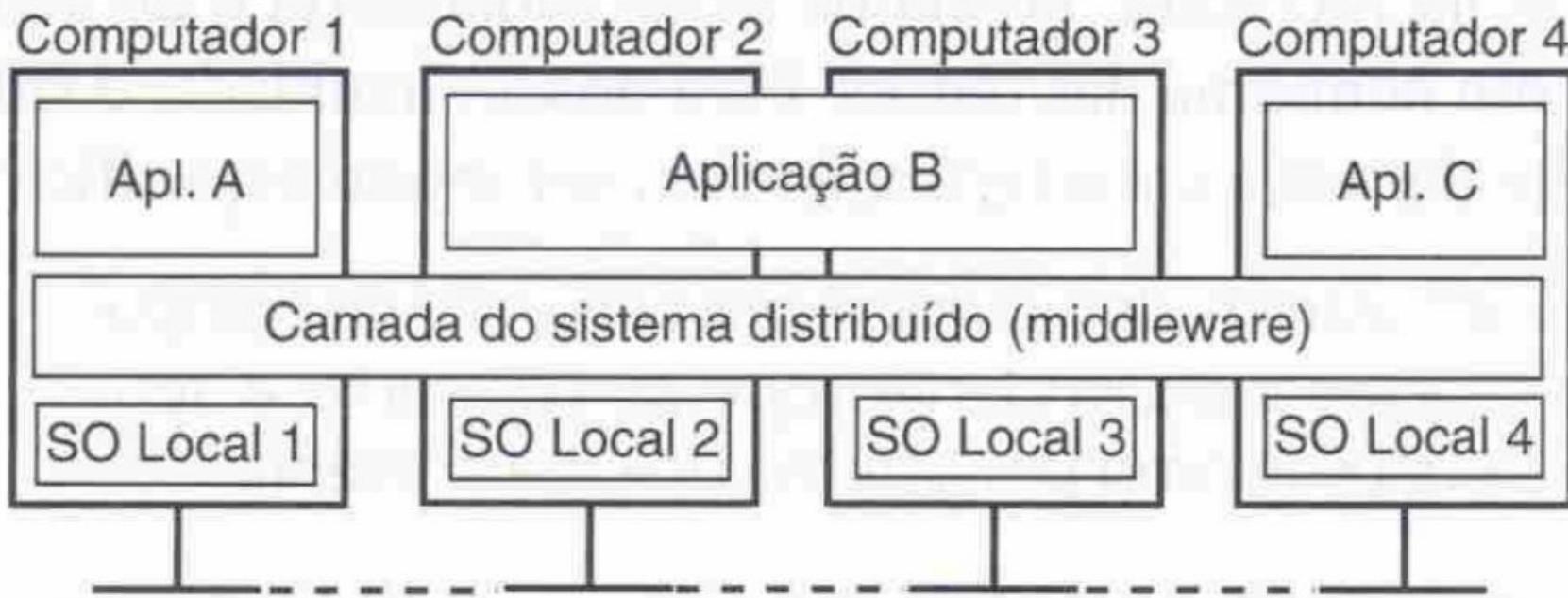
Sistemas Distribuídos Abertos

- Quando adequadamente especificada, uma definição de interface permite que um processo arbitrário que necessite de certa interface se comunique com um outro processo que fornece aquela interface
- Isso permite que duas partes independentes, construam implementações completamente distintas dessas interfaces
- Isso resulta em dois sistemas distribuídos separados que funcionam exatamente da mesma maneira

Middleware

- É o que permite o suporte a redes e sistemas operacionais heterogêneos, formada por todo tipo de computadores oferecendo sempre uma visão única e transparente
- Os sistemas distribuídos costumam ser organizados por meio de uma camada de software
- Essa camada é situada logicamente entre o nível mais alto, composta por usuários e aplicações e uma camada mais baixa, que consiste em sistemas operacionais e facilidades básicas de comunicação

Middleware



Middleware

- Existem diversos tipos de Middleware, como por exemplo:
 - Corba (Common Object Request Broker Architecture)
 - Java RMI
 - Microsoft DCOM
 - XML WebServices (usa protocolo SOAP, que permite a comunicação entre aplicações desenvolvidas em diferentes plataformas e linguagens)
 - RESTful APIs (padrão arquitetural)

Middleware

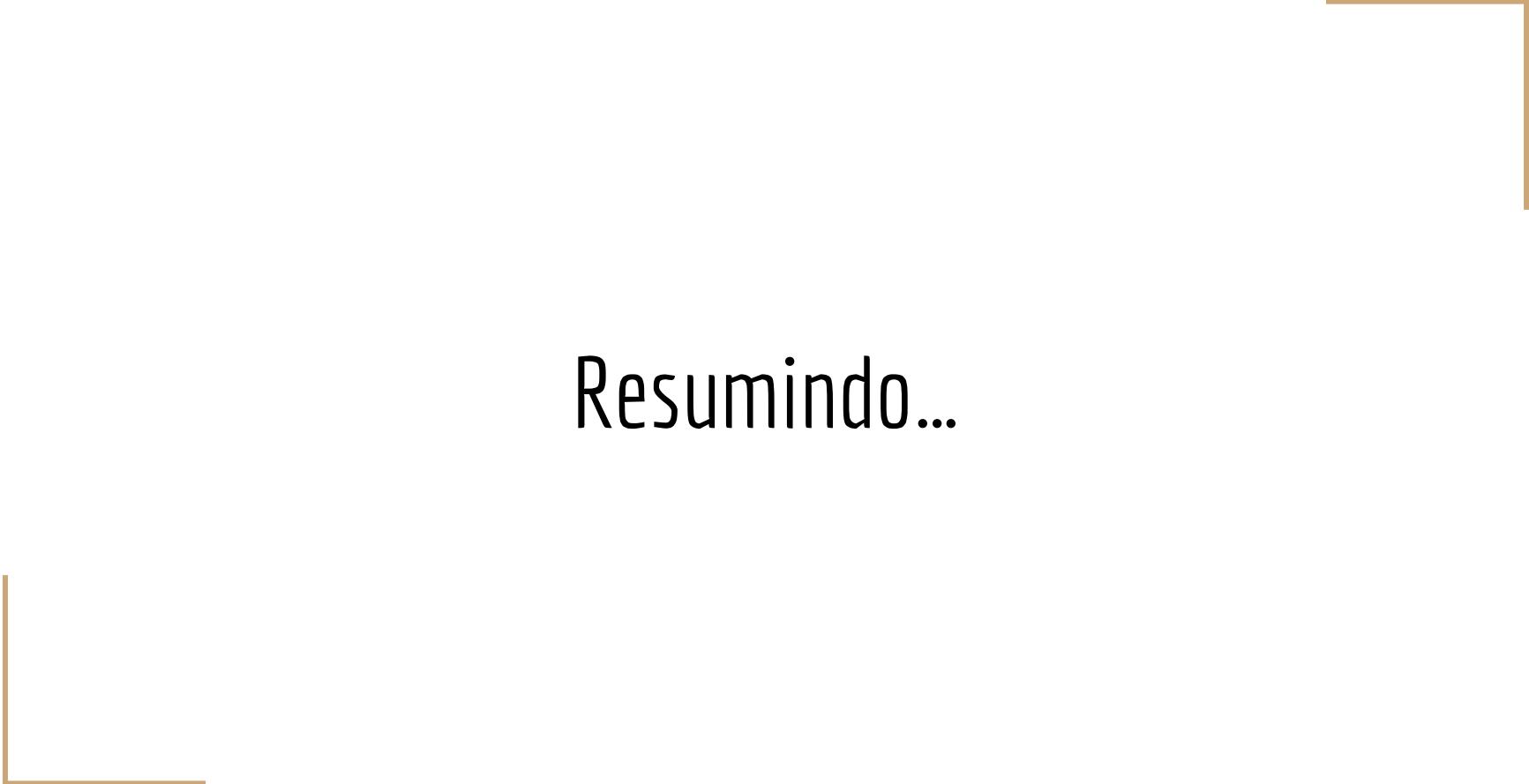
- Laravel 5



Middleware

- **Interoperabilidade**
- A interoperabilidade se caracteriza por determinar até que ponto duas implementações de sistemas ou componentes de fornecedores distintos devem coexistir
- Também dita como trabalhar em conjunto, com base na mera confiança mútua nos serviços de cada um, especificados por um padrão comum (middleware)





Resumindo...

Computação Paralela x Computação Distribuída X Redes de Comunicação

Característica	Redes	S. D.	S. P.
Parece um único processador virtual?	Não	Sim	Sim
Todos tem que rodar um mesmo SO?	Não	Não	Sim
Quantas cópias do SO existem	Várias	Várias	1
Como é feita a comunicação?	Arquivos compartilhados	Troca de Mensagens	Memória compartilhada
É necessária padronização de protocolos?	Sim	Sim	Não
Existe apenas uma fila de tarefas?	Não	Não	Sim
Compartilhamento de arquivos tem semântica bem definida?	Normalmente não	Sim	Sim