

SISTEMAS DISTRIBUÍDOS - NOMEAÇÃO

NOMEAÇÃO

Nomes desempenham papel importante em todos os sistemas de computação

- São usados, por exemplo, para compartilhar recursos e identificar entidades

Para resolver nomes é necessário implementar um Sistema de Nomes

Em um sistema distribuído, a implementação do sistema de nomes costuma ser distribuída por várias máquinas

- A forma como a distribuição é feita influencia diretamente na escalabilidade e eficiência do sistema

NOMEAÇÃO

Modos diferentes e importantes pelos quais nomes são usados em Sistemas Distribuídos:

- Implementação de nomes amigáveis para seres humanos
- Nomes são usados para localizar entidades de modo independente de sua localização no instante em questão
- Entidades podem ser descritas por meio de suas características
 - Mostra formas para resolver um nome de acordo com uma descrição de atributos

NOMES, IDENTIFICADORES E ENDEREÇOS

Nome: cadeia de bits ou caracteres usada para referenciar uma entidade

Entidade: pode ser praticamente qualquer coisa: hosts, impressoras, discos, arquivos, etc

- Entidades são acessadas através de um **ponto de acesso**
- O nome de um ponto de acesso é denominado endereço (exemplo: servidor e seu endereço IP)
- Uma entidade pode fornecer mais de um ponto de acesso
- Um ponto de acesso pode mudar ao longo do tempo

NOMES, IDENTIFICADORES E ENDEREÇOS

O endereço pode ser usado como uma maneira de nomear, identificar uma entidade.

Problema:

- A entidade pode mudar facilmente de ponto de acesso
- Nomes de endereço não são amigáveis para seres humanos

NOMES, IDENTIFICADORES E ENDEREÇOS

Necessário definir um nome que seja independente dos endereços da entidade => endereço independente de localização

- Caso o sistema distribuído seja reorganizado e servidores mudem de endereço é importante que seja possível continuar acessando o serviço

Necessário definir um nome construído para ser usado por seres humanos, normalmente formados por cadeias de caracteres => nome amigável a seres humanos

NOMES, IDENTIFICADORES E ENDEREÇOS

Identificador: um nome que identifica uma entidade e tem as seguintes propriedades:

1. Um identificador referencia, no máximo, uma entidade
2. Cada entidade é referenciada por, no máximo um identificador
3. Um identificador sempre referencia a mesma entidade, isto é, nunca é reutilizado

Uso de identificadores permite que entidades sejam referenciadas sem ambiguidade. Exemplos:

- CPF da pessoa, ISBN de livros, Número de registro do funcionário, Ethernet Addresses, ...

NOMES, IDENTIFICADORES E ENDEREÇOS

- Um sistema de nomeação mantém uma vinculação nome-endereço que em sua forma mais simples é apenas uma tabela de pares (nome, endereço)
- Em sistemas distribuídos de grande escala é necessário usar tabelas descentralizadas
- Um nome é decomposto em várias partes e a resolução é feita por meio de consultas recursivas das partes:

<https://www.fho.edu.br>

NOMEAÇÃO SIMPLES

- Identificadores são, em muitos casos, cadeias aleatórias de bits (nomes não estruturados ou simples);
- Um identificador não contém informação de como localizar uma entidade
 - Como nomes simples podem ser resolvidos?
- Soluções simples aplicáveis somente em redes locais:
 - Broadcasting e Multicasting
 - Ponteiros repassadores

Possuem problemas de escalabilidade!

NOMEAÇÃO SIMPLES - BROADCASTING

Broadcasting

Recursos oferecidos a redes locais nas quais todas as máquinas estão conectadas a um único cabo ou seu equivalente lógico

Como funciona:

- Uma mensagem que contém o identificador da entidade é enviada a cada máquina da rede
- Cada uma delas deve verificar se tem essa entidade
- Máquinas com ponto de acesso para a entidade, enviam uma mensagem que contém o endereço

NOMEAÇÃO SIMPLES - BROADCASTING

Broadcasting

Se torna ineficiente quando a rede cresce

- Largura de banda da rede é desperdiçada, com grande número de mensagens de requisição
- Aumento da probabilidade de colisões de mensagens, diminuindo a quantidade de dados transferidos do sistema
- Grande número de máquinas pode ser interrompido por requisições que não podem responder

NOMEAÇÃO SIMPLES - MULTICASTING

Multicasting

- Grupos restritos de hosts são identificados por um endereço multicast
- Quando um host envia uma mensagem a um endereço multicast, a camada de rede fornece um serviço para entregar aquela requisição a todos os membros do grupo
- Envia apenas para grupos específicos, limitando o tráfego
- Também usado para associar réplicas e localizar a réplica mais próxima
- A réplica mais próxima é aquela cuja a resposta chega primeiro

NOMEAÇÃO SIMPLES - PONTEIROS REPASSADORES

Ponteiros Repassadores

Usado para localização de entidades móveis.

O princípio é bem simples:

- Quando uma entidade se move de A para B, deixa em A uma referência à sua nova localização, em B.

Tão logo uma entidade é localizada, o cliente pode consultar o endereço corrente da entidade percorrendo uma cadeia de ponteiros repassadores.

NOMEAÇÃO SIMPLES - PONTEIROS REPASSADORES

Ponteiros Repassadores

Se não houver precauções, a cadeia para uma entidade de alta mobilidade torna o custo de localização proibitivo.

Todas localizações intermediárias terão de manter sua parte da cadeia de ponteiros repassadores pelo tempo necessário.

Enlaces rompidos geram vulnerabilidade

- Se um ponteiro repassador for perdido, a entidade não poderá ser alcançada

É importante manter cadeias curtas e repassadores robustos.

NOMEAÇÃO SIMPLES - PONTEIROS REPASSADORES

Ponteiros Repassadores

Se não houver precauções, a cadeia para uma entidade de alta mobilidade torna o custo de localização proibitivo.

Todas localizações intermediárias terão de manter sua parte da cadeia de ponteiros repassadores pelo tempo necessário.

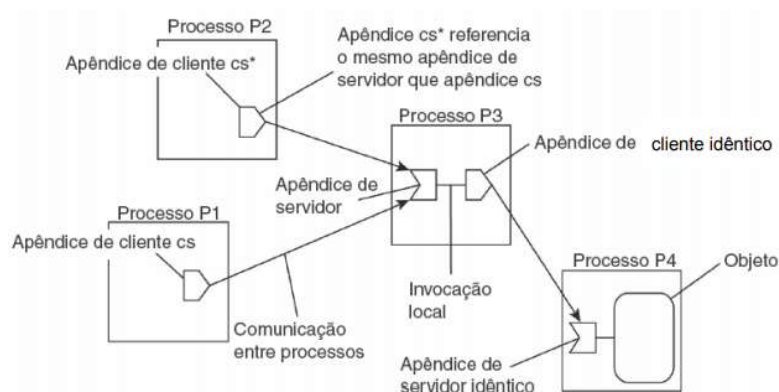
Enlaces rompidos geram vulnerabilidade

- Se um ponteiro repassador for perdido, a entidade não poderá ser alcançada

É importante manter cadeias curtas e repassadores robustos.

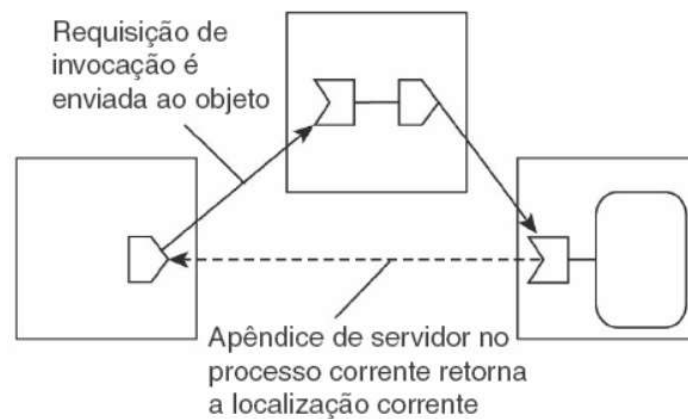
NOMEAÇÃO SIMPLES - PONTEIROS REPASSADORES

Ponteiros Repassadores - Princípio



NOMEAÇÃO SIMPLES - PONTEIROS REPASSADORES

Ponteiros Repassadores - Redirecionamento



NOMEAÇÃO SIMPLES - LOCALIZAÇÃO NATIVA

Localização Nativa

Quando uma entidade móvel se torna inalcançável por algum motivo
=> uma solução é a localização nativa

Localização Nativa

- A máquina em que o objeto foi criado mantém sempre uma referência da sua localização corrente
- Esta referência é mantida de modo tolerante a falha
- Quando uma cadeia é rompida, pergunta-se à localização nativa do objeto qual a sua localização corrente

NOMEAÇÃO SIMPLES - LOCALIZAÇÃO NATIVA

Localização Nativa

- Uma localização nativa monitora a localização corrente de uma entidade
- Usada como mecanismo de emergência para localização de serviços baseada em ponteiro de repassadores

Exemplo:

- Funcionamento do Mobile IP (mIPv6)
 - Cada host móvel usa um endereço IP fixo => Localização Nativa

NOMEAÇÃO SIMPLES - DHT

Tabelas de Hash Distribuídas (DHT)

Como resolver com eficiência uma chave k para o endereço de $\text{succ}(k)$?

- Uma abordagem simples porém não escalável => abordagem linear
- Permite que cada nó p monitore o sucessor e o seu predecessor
- Quando p receber uma requisição para a chave k :
 - Se $\text{pred}(p) < k \leq p$, então p retorna o próprio endereço
 - Senão p repassa a requisição para os seus vizinhos

NOMEAÇÃO SIMPLES - DHT

Tabelas de Hash Distribuídas (DHT)

Será considerado o sistema Chord => nós são organizados logicamente em um anel

- Usa um espaço de endereçamento de m bits para chaves para designar nós e entidades específicas (arquivos, processos)
- Uma entidade com chave k cai sob a jurisdição do nó que tenha o menor identificador $id \geq k$, nó denominado sucessor de $k \Rightarrow succ(k)$

NOMEAÇÃO SIMPLES - DHT

Tabelas de Hash Distribuídas (DHT)

Ao invés desta abordagem linear, considere que cada nó Chord mantenha uma Tabela de Derivação

- A tabela de derivação do nó p é denotada por FT_p
- A tabela possui, no máximo, m entradas FT

$$p[i] = succ(p + 2^i - 1)$$

- A ideia é que uma consulta possa ser resolvida em $O(\log N)$ passos

NOMEAÇÃO SIMPLES - DHT

Tabelas de Hash Distribuídas (DHT)

Aplicações:

- [BitTorrent](#): Distribuição de arquivos. BitTorrent usa opcionalmente uma DHT como um tracker distribuído para prover um ponto de encontro entre os clientes fazendo download de um arquivo particular.
- [The Circle](#): Compartilhamento de arquivos e chat
- [eMule](#): Compartilhamento de arquivos
- [JXTA](#): Plataforma peer-to-peer de código aberto

NOMEAÇÃO ESTRUTURADA

- Nomes simples são bons para as máquinas, mas não são convenientes para seres humanos
 - Uma alternativa é usar nomeação estruturada
- Sistemas de nomeação geralmente suportam nomes estruturados que são compostos por nomes simples
 - Usado para facilitar a interface com humanos
- A nomeação de arquivos e de hosts na Internet segue esta abordagem
- Nomes são organizados em espaços de nomes

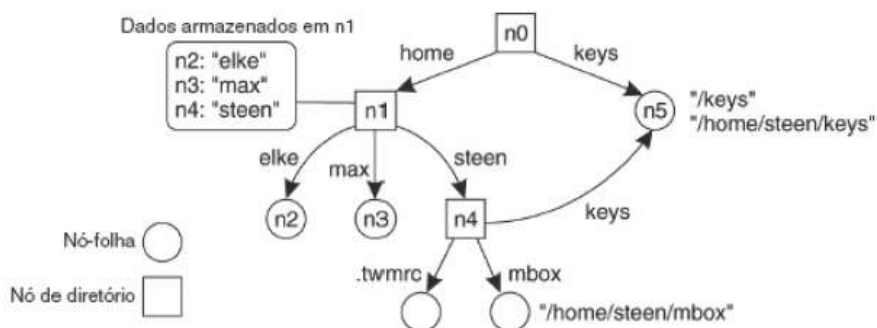
NOMEAÇÃO ESTRUTURADA

Espaço de nomes para nomes estruturados podem ser representados por um grafo direcionado rotulado com dois tipos de nós:

- Nó-folha: entidade nomeada, não tem ramos de saída
 - Armazena informações sobre a entidade que está representando => por exemplo, seu endereço
- Nó de diretório: possui vários ramos de saída, cada um rotulado com um nome
 - Armazena tabelas de diretório que é um par => rótulo do ramo, identificador do nó
- Um nó que não tem ramos de entrada é denominado raiz

NOMEAÇÃO ESTRUTURADA

Grafo de Nomeação



NOMEAÇÃO ESTRUTURADA

- Espaços de nomes oferecem um mecanismo para armazenar e recuperar informações sobre entidades por meio dos seus nomes
- Dado o nome de um caminho deve ser possível consultar qualquer informação no nó referenciado por aquele caminho
- O processo de busca de um nome é denominado => resolução de nomes

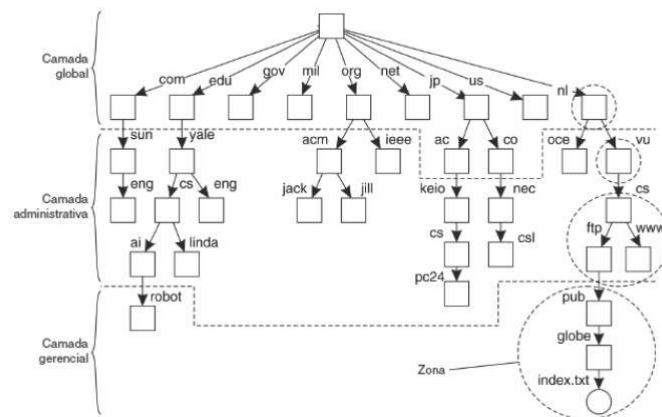
/home/steen/mbox

DNS - DOMAIN NAME SYSTEM

- Permite que hosts, roteadores, servidores de nomes se comuniquem para resolver nomes
- Função imprescindível da Internet implementada como protocolo de camada de aplicação
- Base de dados distribuída implementada na hierarquia de muitos servidores de nomes
 - Uma base de dados centralizada não seria escalável => Um ponto único de falha, grande volume de tráfego, base de dados centralizada e distante

DNS - DOMAIN NAME SYSTEM

Em essência o espaço de nomes DNS pode ser dividido em uma camada global e uma camada administrativa



COMPLEMENTO

Sistemas Distribuídos - Sistemas de Nomeação

<https://www.youtube.com/watch?v=YEl6qlenqqE>

Sistemas Distribuídos - Abordagens de Nomeação

<https://www.youtube.com/watch?v=xLRPNW0mGFc>