

Sistemas Distribuídos e Tolerância a Falhas

Comunicação em Grupo

Trabalho realizado por:
1486 Etelvina Pinho
1489 Edgar Pinto

Objectivo da comunicação em grupo

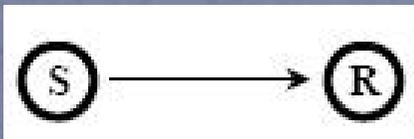
- Resolver problemas de inconsistências na comunicação entre processos distribuídos que cooperam na execução de uma tarefa
- Garantindo a correcção da comunicação apesar da ocorrência de falhas

Aplicações da comunicação em grupo

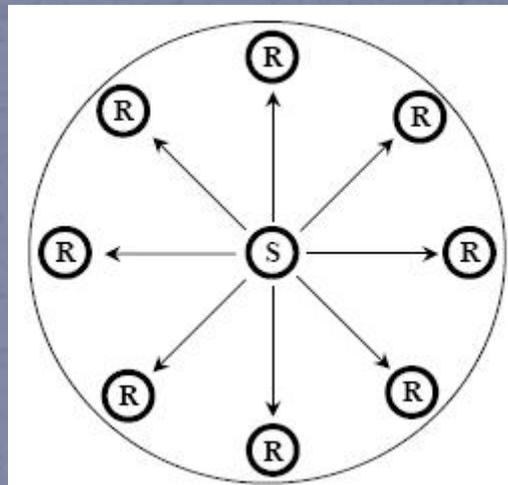
- Na difusão em grande escala
- Nos sistemas tolerantes a falhas

O que é um grupo

- Conjunto de processos que cooperam
- A mensagem é enviada para todos os membros do grupo



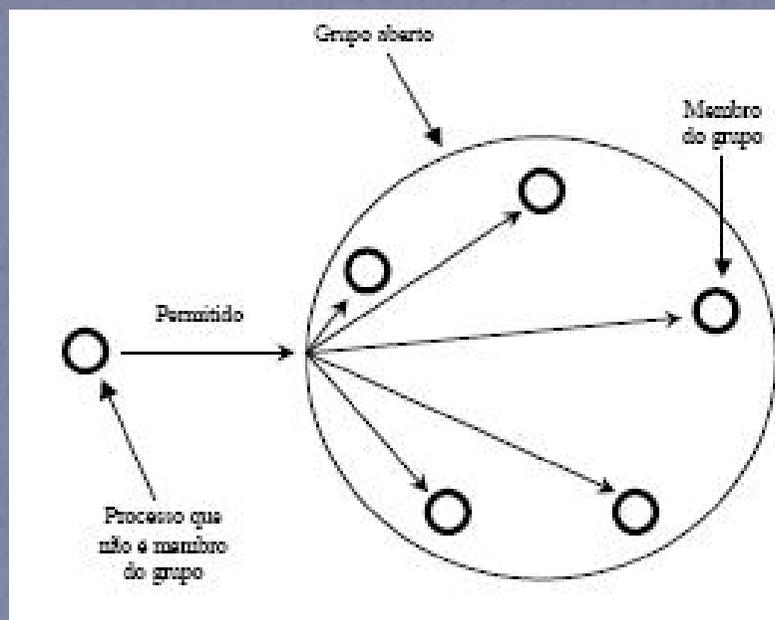
Comunicação
“ponto a ponto”



Comunicação “um para
muitos”

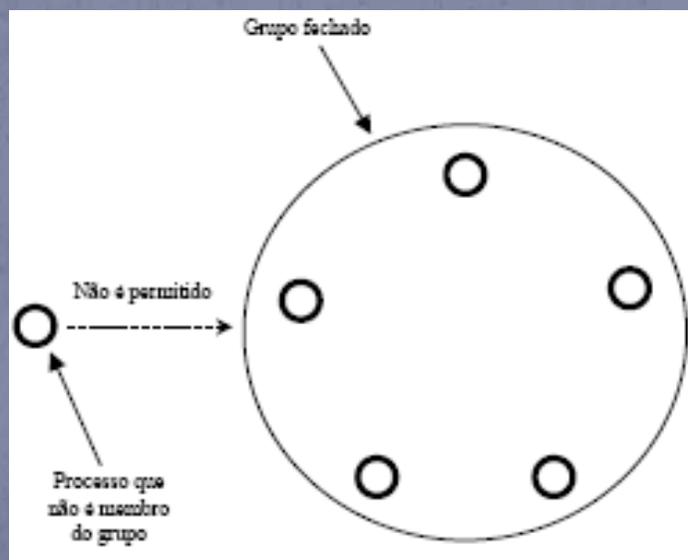
Grupos abertos

- Grupos abertos
 - Qualquer processo no sistema pode enviar mensagens para o grupo
 - Usados tipicamente nos casos de replicação de serviços



Grupos fechados

- Grupos fechados
 - Só os membros do grupo podem enviar mensagens para o grupo
 - Processos fora do grupo não podem enviar mensagens ao grupo como um todo.
 - Apenas a membros individualmente
 - Usados tipicamente em processamento paralelo



Grupos Visíveis e Invisíveis

- Visíveis
 - Existe informação actualizada sobre os membros que pertencem ao grupo
- Invisíveis
 - Um grupo refere-se a um conjunto de réplicas do mesmo componente que executa a mesma acção
 - Replicação é uma técnica de tolerância falhas que assegura a continuidade dos serviços apesar da falha de componentes individuais

Grupos dinâmicos!

Porquê?

- Criar e destruir grupos
- Processos podem entrar ou deixar um grupo
- Um processo pode pertencer a mais que um grupo

Necessário fazer em grupos dinâmicos

- Controlar quais os membros do grupo
- Permitir a inclusão e exclusão de membros

Controlar os membros do grupo

- Servidor de grupo
 - Todos os pedidos são feitos ao servidor
 - Deve manter uma base de dados completa
 - Eficiente
 - Fácil de implementar
- Controle de forma distribuída
 - Cada novo membro deve enviar uma mensagem a anunciar a sua presença
 - Ao sair do grupo deve avisar todos os membros do grupo do qual vai sair
 - Facilita uma implementação tolerante a falhas
 - A entrada/saída do grupo deve ser sincronizadas com troca de mensagens

Grupos e Vistas

- Serviços de membros de grupo
 - Propriedades,
 - Exactidão: a informação fornecida reflecte o cenário físico
 - Coerência: a informação fornecida é coerente em todos os processos
 - Proporciona,
 - Habilidade para criar e se tornar membro de um grupo
 - Disposição de informação sobre quais os membros que estão acessíveis
 - “Group view”: os processos têm um único identificador estão ordenados pelo mesmo

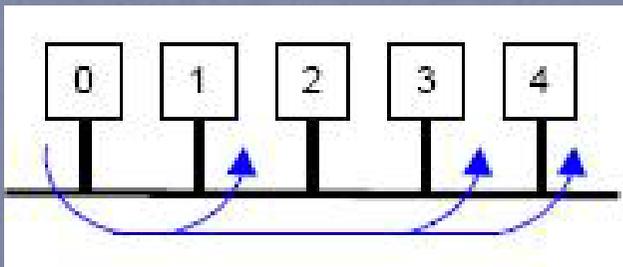
Grupos e Vistas

- Serviços de membros de grupo
 - Garante,
 - Aspectos de fiabilidade: garante a entrega de mensagens
 - Aspectos de ordenação: garante a ordenação de mensagens
- Funcionamento do “group views”
 - Quando um membro entra/sai de um grupo
 - Uma nova vista é enviada a todos os membros do grupo
 - É garantida a ordem de entrega da mensagem de vista

Implementação da comunicação em grupo

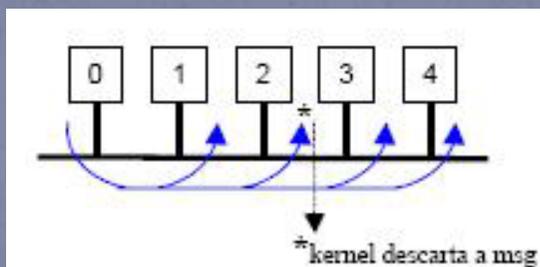
- Multicast

- É enviada uma mensagem apenas às máquinas do grupo
- Múltiplas máquinas podem ouvir



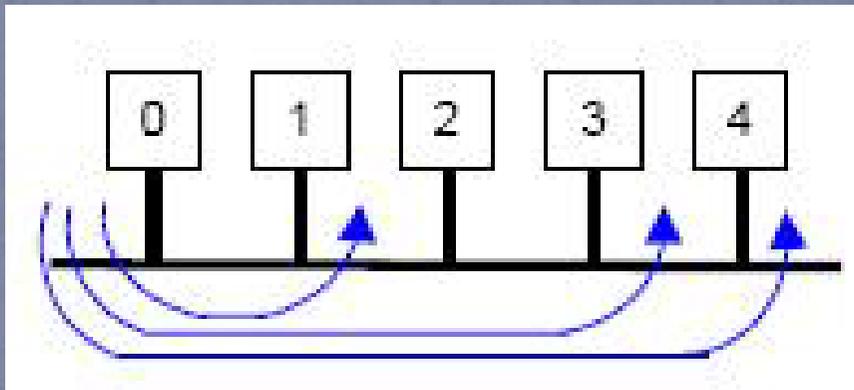
- Broadcast

- É enviada uma mensagem a todas as máquinas
- Todas as máquinas podem ouvir



Implementação da comunicação em grupo

- Unicast
 - São enviadas n mensagens (n envios)
 - Uma para cada membro do grupo



Tipos endereçamento

Tipo	IPv4	IPv6	TCP	UDP	Destinatários	Receptores
unicast	Sim	Sim	Sim	Sim	1	1
anycast	Não	Sim...	?	Sim	Grupo	1
multicast	Opc.	Sim	Não	Sim	Grupo	Grupo
broadcast	Sim	Não	Não	Sim	Todos	Todos

Protocolo Multicast

- Componentes principais:
 - Endereçamento
 - Selecciona o caminho da mensagem desde a origem aos destinos
 - Encontra o caminho que minimiza tanto o número de mensagens trocadas com a latência de multicast
 - Tolerância na omissão
 - Lida com as mensagens que são perdidas ou corrompidas na infra-estrutura física, através da transmissão redundante ou retransmissão
 - As omissões são toleradas usando acknowledgments para detectar erros e retransmitir mensagens perdidas

Protocolo Multicast

- Controlo de fluxo
 - Minimiza a perda de dados provocada pela falta de espaço de buffer nos endereços
- Ordem
- Recuperação de falhas
 - Impõe uma ordem predefinida e um critério de fiabilidade em relação às mudanças de vista
- Os métodos apresentados minimizam o tráfego de rede e oferecem uma detecção mais rápida de falhas

Endereços de Multicast

- IP v4: 224.0.0.0 a 239.255.255.255 (class D):
 - 32-bit: endereço do grupo
 - 28-bit: ID do grupo
 - Endereços reservados (administração):
 - 224.0.0.0 – 224.0.0.255
- IP v6: 0xff...
 - 128-bit: endereço do grupo
 - 112-bit: ID do grupo

Propriedades da comunicação em grupo

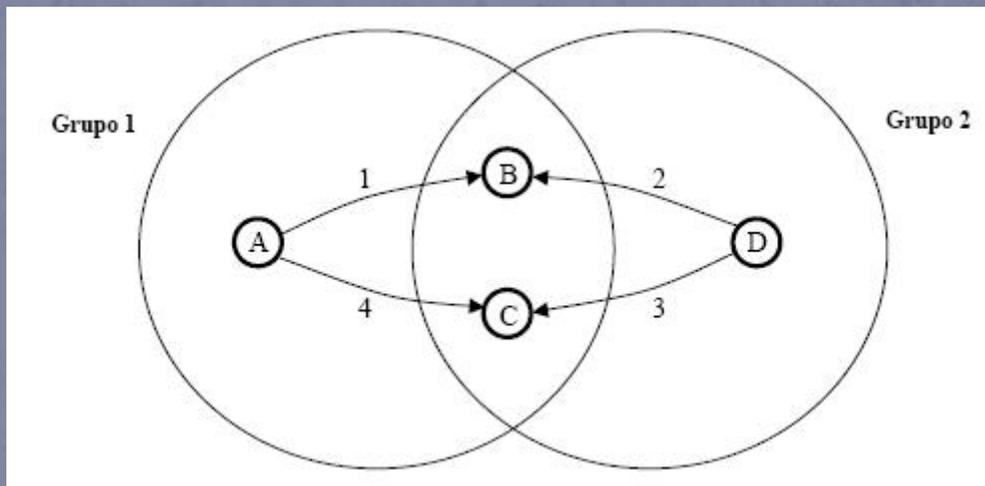
- Atomicidade
 - Propriedade de entregar “tudo ou nada”
 - Entrega de mensagens a todos os membros do grupo
 - Torna a programação mais fácil
 - Quando um processo envia uma mensagem para um grupo não precisa se preocupar com o facto de alguém não ter recebido a mensagem
 - Falhas na entrega são comunicadas ao emissor
 - Pode realizar as acções necessárias para a recuperação
 - A única forma de garantir a entrega da mensagem para todos os destinos é através do envio de mensagens de reconhecimento

Propriedades da comunicação em grupo

- Ordenação de mensagens
 - Ordem pela qual as mensagens são entregues
- Vantagem:
 - Torna a comunicação em grupos mais simples
- Desvantagem
 - Acontece se, mais que um processo envia uma mensagem para o grupo ao mesmo tempo
 - Logo, não é possível determinar a ordem pela qual as mensagens chegam

Propriedades da comunicação em grupo

- Sobreposição de grupos
 - Processos podem pertencer a mais que um grupo
 - Podem surgir inconsistências



Propriedades da comunicação em grupo

- Escalabilidade
 - Problemas:
 - Multicast: necessidade de algoritmos mais sofisticados mantendo rastros
 - O uso de gateways dificulta uma ordenação por tempo global absoluto: mais de um pacote ao mesmo tempo na rede
 - Algoritmos escaláveis: aumento da complexidade

Primitivas de comunicação

- De modo a garantir que uma mensagem seja entregue a todos os membros do grupo
 - Requerer-se um acknowledged de cada membro
 - Se algum acknowledged está desaparecido o emissor retransmite a mensagem
 - Para prevenir a retransmissão consecutiva da mensagem no caso de alguns membros falharem é requerido o mecanismo detector de falhas que indica se deve ou não esperar por uma resposta desse membro

Primitivas de comunicação

- Quando membros de um grupo fornecem um serviço a clientes que não pertencem ao grupo:
 - Os clientes enviam um pedido multicast ao grupo e esperam uma resposta de um ou todos os membros do grupo
 - O processamento desse pedido pode ser distribuído pelos membros (Load Balancing)
- O emissor faz um pedido multicast a todos os membros, está interessado em receber apenas uma única resposta, para isso bloqueia até que a primeira resposta seja obtida

Primitivas de comunicação

- Os grupos podem ser usados para paralelizar trabalho
 - Cada processador é responsável por executar uma parte da tarefa
 - O cliente envia um pedido para o grupo mas bloqueia até que todas as respostas cheguem

Dificuldades da comunicação em grupo

- Grupos numerosos geograficamente distribuídos
- Dificuldade:
 - Elevada latência na recolha de confirmações
 - Heterogeneidade dos elementos do grupo
 - As mensagens necessitam ser armazenadas durante longos períodos de tempo para eventual retransmissão
 - Pode levar a um eventual congestionamento de todo o sistema

Conclusão

- A comunicação em grupo, através do protocolo multicast, garante a correção da comunicação ao assegurar a entrega de mensagens e a sua ordenação, apesar da ocorrência de falhas. No entanto ainda é de difícil implementação em grupos numerosos.

Referências

- <http://www.lasid.ufba.br/projetos/tolerancia.html>
- http://www.navigators.di.fc.ul.pt/archive/jorge_concretiz.pdf
- <http://web.fe.up.pt/~jmcruz/sd/acetat/2-comun-multicast.pdf>
- <http://pdos.csail.mit.edu/~kaashoek/phdthesis>
- <http://www.gsd.inesc-id.pt/~ler/docencia/tfd0506/artigos-v1/18.pdf>
- <http://www.gsd.inesc-id.pt/~ler/docencia/tfd0607/bib/appia-group.pdf>
- <http://www.frb.br/ciente/2005.2/BSI/BSI.LIMA&.F3.pdf>