

Complementos de Bases de Dados – Replicação – (em MS SS)

Engenharia Informática
2º Ano / 1º Semestre

Cláudio Miguel Sapateiro
claudio.sapateiro@estsetubal.ips.pt

Sumário

- Motivação e Objetivos
- Cenários
- Metáfora *Publisher-Subscriber*
- Tipos de Replicação e Subscrição
- Fatores a considerar
- Agentes envolvidos

Replicação

Motivação

- Numa arquitetura completamente centralizada, se
 - Muitos acessos (concorrentes)
 - Muitos dados persistidos
 - Muitos dados trocados

Podemos ter problemas de:

- Performance
- Disponibilidade
- Manutenção

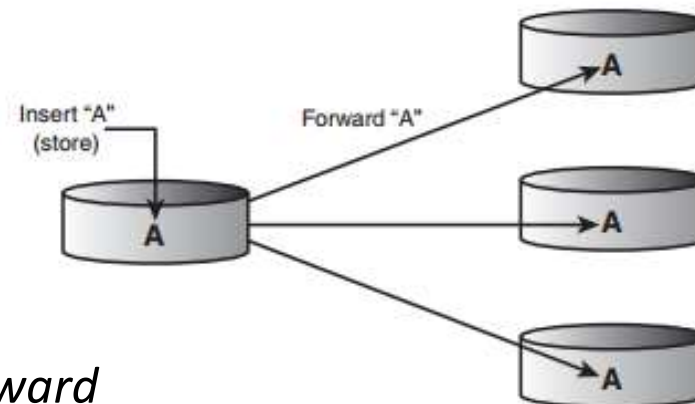
❖ Uma das opções a considerar passa pelo recurso à **Replicação**

Replicação

Objetivos

“Replication is a set of technologies for copying and distributing data and database objects from one database to another and then synchronizing between databases to maintain consistency.”

- A replicação permite manter múltiplas copias de partes de uma BD em diferentes/multiplos locais



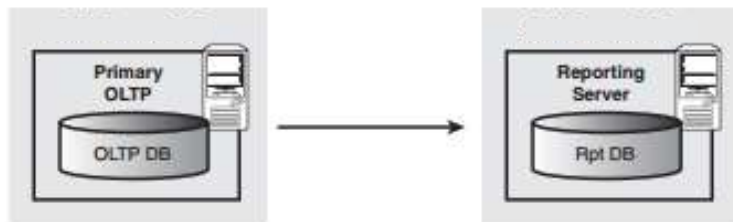
Modelo clássico:
store-and-forward

Replicação

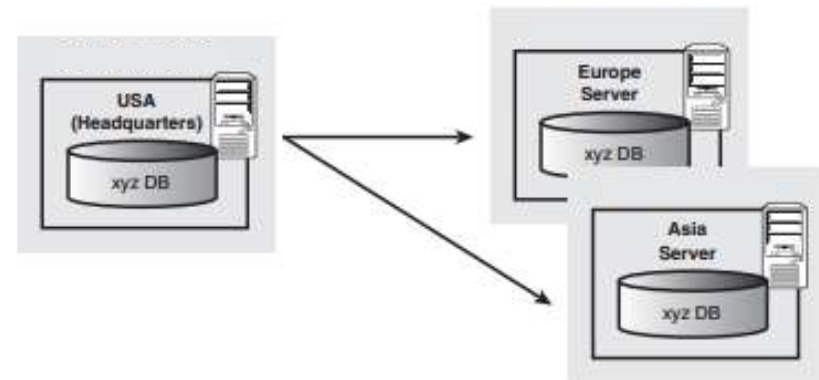
Variantes

Cenários e.g.

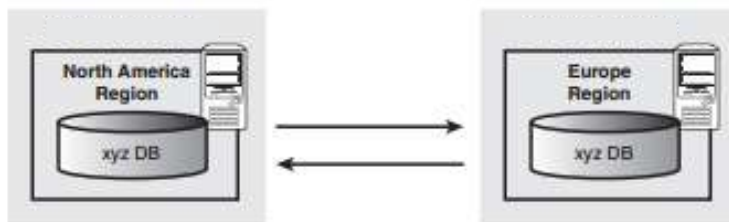
isolate reporting



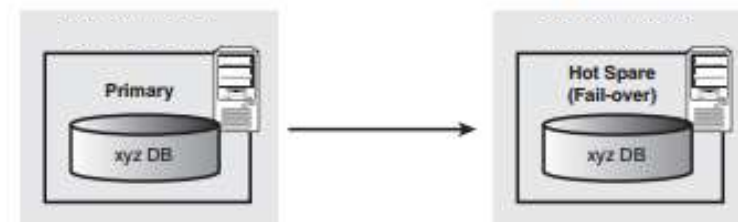
partitioning by region



bidiretional



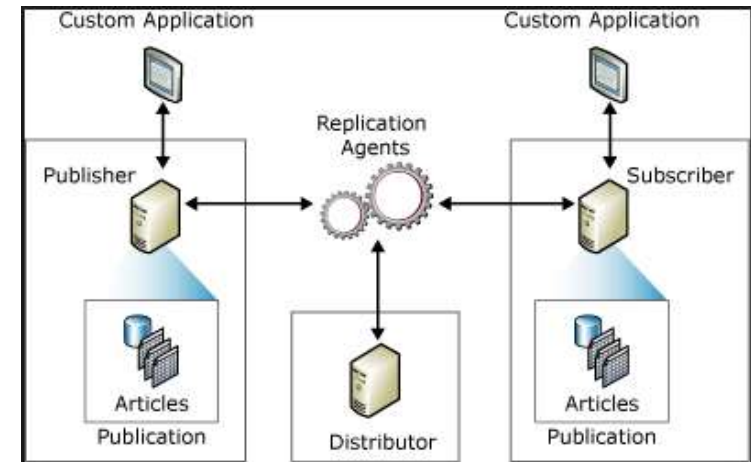
failover server



Replicação

Metáfora *Publisher-Subscriber*

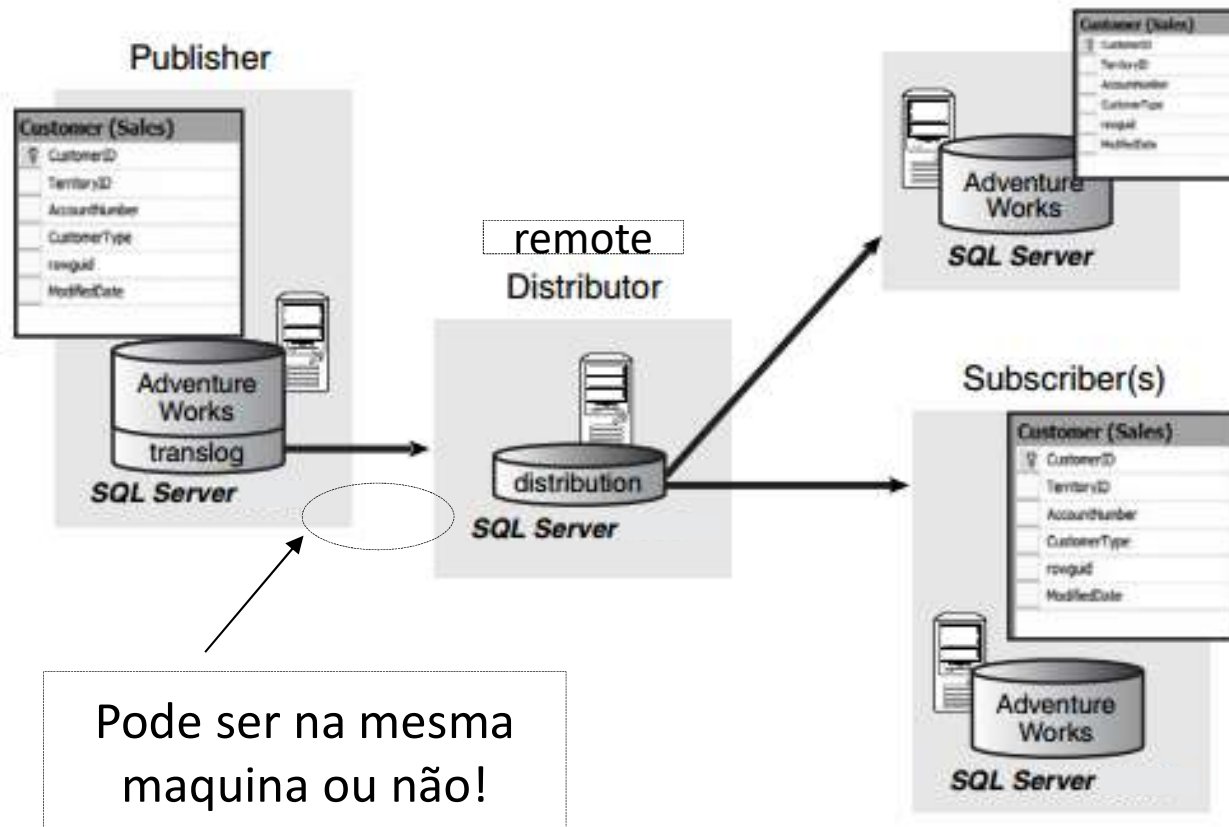
- ***Distributor***: servidor de BD responsável pela coordenação da operação de replicação
 - Armazena a base de dados *distribution* (*store-and-forward*) (contém a metadata, dados históricos e transações).
 - Tem um papel mais relevante nas replicações do tipo *snapshot* e *transaccional*
- ***Publisher***: servidor de BD responsável pela manutenção da informação na BD de origem e envio da informação alterada, para o ***Distributor***
- ***Subscriber***: servidor de BD responsável pela manutenção da cópia da informação e pela receção da informação oriunda do ***Distributor***



Replicação

Metáfora *Publisher-Subscriber*

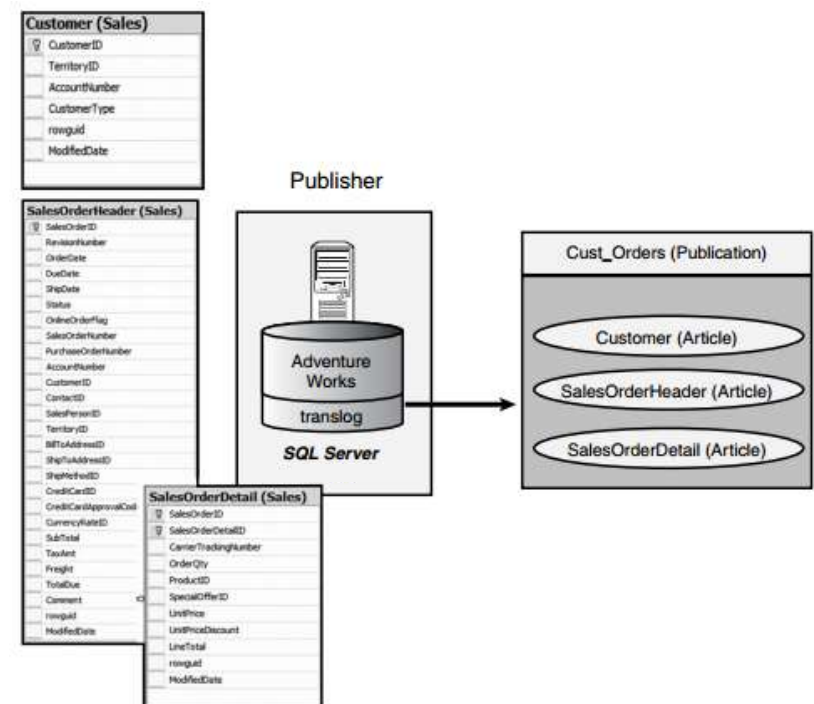
Exemplo



Replicação

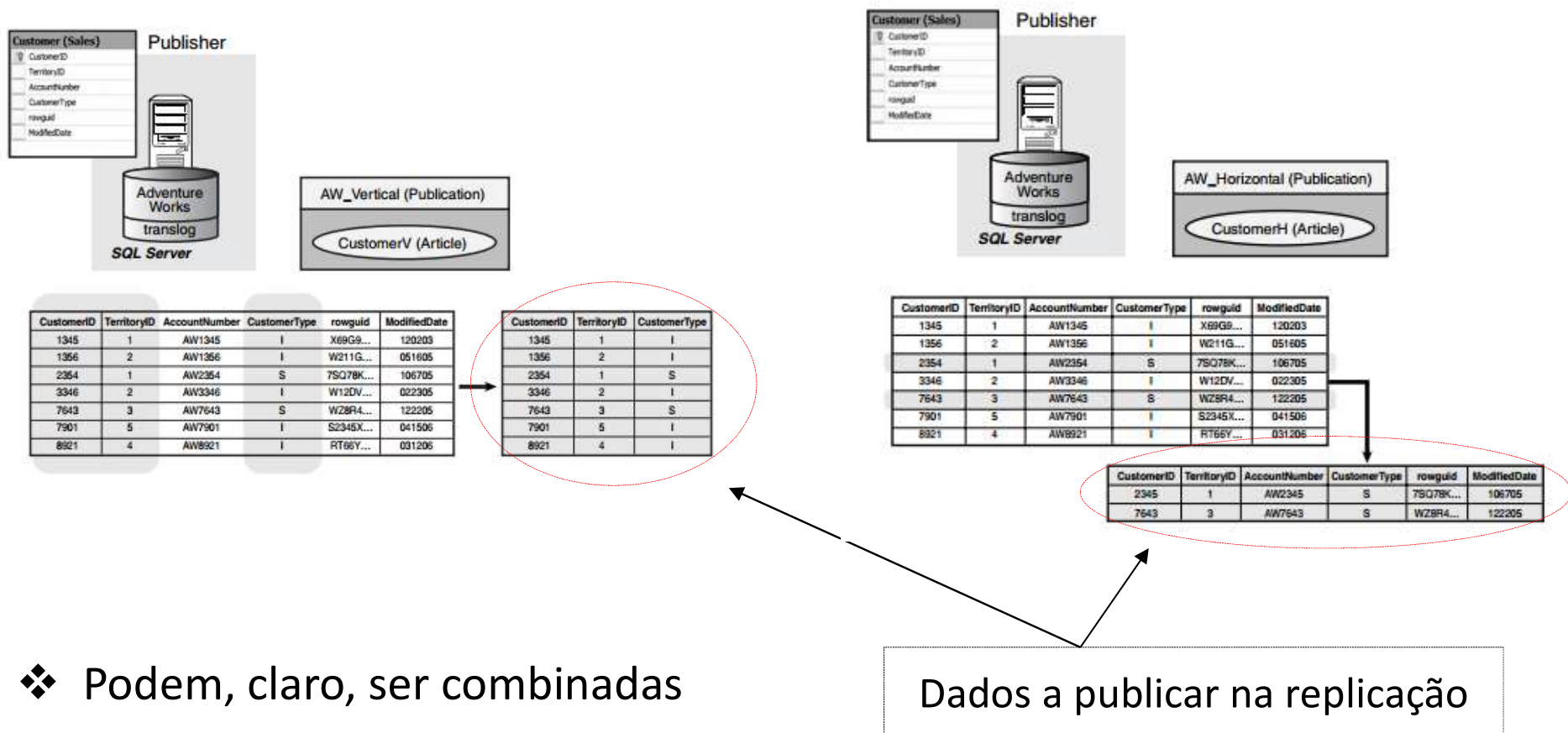
Informação a Replicar

- Artigo:
 - Dados (toda ou, parte da informação, de uma tabela) ou
 - Objectos da BD (stored procedures, views, índices, funções definidas pelo utilizador);
- Publicação:
 - conjunto de um ou mais artigos, que constitui uma unidade de subscrição
- ❖ As réplicas podem ser referentes a 1 ou mais publicações



Replicação

Filtragem de Artigos



Replicação

Tipos de Replicação

- Snapshot
 - Transacional
 - Merge
- ❖ Um tipo de replicação aplica-se a uma publicação
- Assim podem coexistir vários tipos de replicação na mesma BD

Replicação

Snapshot

Quando ocorre o momento da sincronização, é gerada e enviada aos *Subscribers* uma fotografia dos dados, i.e.,

- ✓ são replicados os dados exatamente como se encontram em determinado momento, sem monitorização das alterações

- Quando Utilizar:
 - Com alterações substanciais mas pouco frequentes dos dados
 - Os *Subscribers* necessitam de acesso a dados apenas em leitura
 - Existe um grande desfasamento nos períodos de atualização dos dados

Replicação

Transacional

Quando as alterações incrementais na base de dados de origem são replicadas na base de dados de destino através do *Distributor*, minimizando o período de desfasamento da informação entre bases de dados

- ✓ os dados alterados nas bases de dados envolvidas, são atualizados consistentemente, i.e., assegurada a consistência transacional

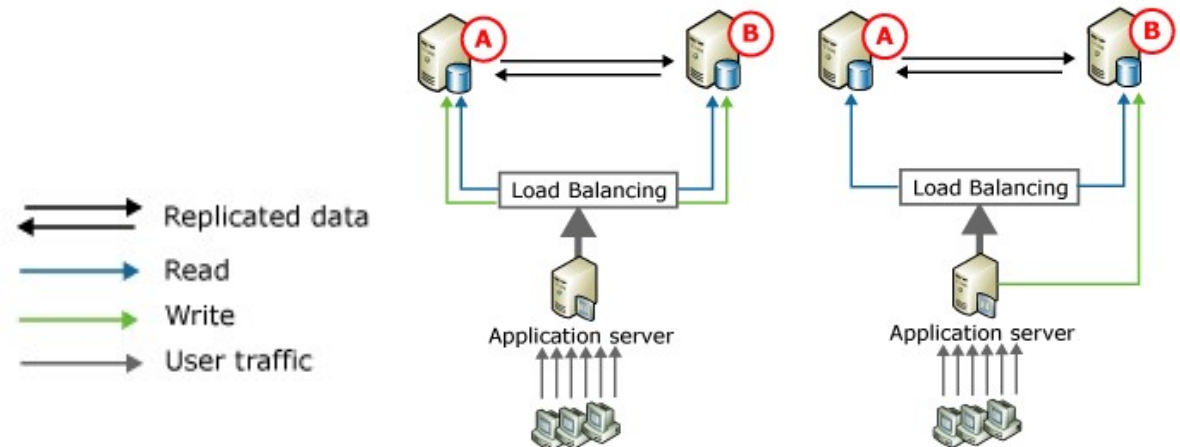
- Quando Utilizar:
 - Os *Subscribers* devem receber as alterações com a maior brevidade possível e refletindo transações originais;
 - Maioritariamente em ***server-to-server***

Replicação

Transacional (cont.)

Outros Tipos

- Disponível também numa versão de ***updtale subscriptions***
Permite atualizações de dados nos subscribers (contudo pouco frequentes) com envio para o publisher (e propagada aos restantes subscribers)
- **Peer-to-peer:**
 - alta disponibilidade, escalabilidade e desempenho
 - Múltiplos servidores
- ***Bidirectional***
 - Só entre 2 servidores simultaneamente publishers/subscribers



Replicação

Merge

Caso em que é permitido que as bases de dados – *subscriber* – efetuem alterações autonomamente, procedendo-se posteriormente (periodicamente ou a pedido) à reconciliação das alterações

- Quando Utilizar:
 - Cenários **Server-to-Client**
 - Os *Subscribers* necessitam de efectuar alterações e propagá-las no *Publisher*
 - Os *Subscribers* necessitam de receber dados efectuar alterações *offline* e sincronizá-las no *Publisher* e nos outros *Subscribers*;
 - *Subscribers* trabalham com uma parte da publicação e não com a publicação total (e.g. via filtragem)

Replicação

Formas de Subscrição

- ***Push Subscription:***
a subscrição é definida para um ou vários *Subscribers*, no *Publisher* aquando da alteração da Publicação.
As alterações são assim enviadas para os *Subscribers* logo que ocorrem no *Publisher*
 - e.g. cenários de múltiplos *subscribers* com um *distributor* dedicado (recursos alocados para o efeito)
- ***Pull Subscription:***
a subscrição é definida e iniciada no *Subscriber*, para as Publicações disponibilizadas pelo *Publisher*
(acessíveis aos *Subscribers* que se registem para o efeito)
 - e.g. se o *publisher* e o *distributor* estão na mesma maquina, pode ser pertinente avaliar a execução dos agentes/serviços responsáveis pela replicação nos *subscribers* para aliviar a maquina *publisher*

Replicação

Fatores a Considerar

- Na seleção do tipo de replicação de deve atender-se, a
 - Grau de Desfasamento Aceitável
 - Grau de Autonomia da(s) Base de Dados
 - Nível de Consistência Transacional
 - Numero de nós, e o seu papel
 - Objectivos/requisitos no balanceamento de carga

Replicação

Agentes

O MS SS utiliza diversos agentes para conduzir as tarefas associadas ao processo de replicação

- ***Snapshot:***
prepara o esquema da base de dados, as tabelas publicadas e as *stored procedures* necessárias à replicação,
 - armazenando essa informação na base de dados *distribution*
 - Corre no *Distributor*
- ***Distribution:***
utilizado na replicação *Snapshot* e *Transaccional*, por forma a implementar a definição da replicação estabelecida na base de dados *distribution* no *Distributor*
 - Corre no *Distributor* (*Push Subscription*) e nos *Subscribers* (*Pull Subscription*)

Replicação

Agentes II

- **Log Reader:** utilizado na replicação Transaccional
 - copia todas as transacções marcadas para replicação, do log de transações do Publisher, para a base de dados distribution no Distributor, para que este as possa distribuir pelos Subscribers.
 - corre em cada uma das bases de dados.
- **Merge:** utilizado na replicação do tipo Merge
 - aplica um snapshot inicial ao *Subscriber* para, com esta base, proceder futuramente à reconciliação das alterações que venham a ocorrer, do *Subscriber* para o *Publisher* e vice-versa.
 - corre no Distributor (Push Subscription) e nos Subscribers (Pull Subscription);
- **Queue Reader:** utilizado na replicação Transaccional,
 - responsável pelo tratamento das mensagens em fila e sua aplicação às respectivas Publicações.
 - corre no *Distributor*;

mini Sumário

1. Cenários de replicação
2. Conceito replicação, artigo e publicação
3. Componentes da replicação
4. Tipos de replicação

10:00



Exercícios

1. Qual a motivação para uma replicação num cenário *“isolate reporting”*?
2. Qual tem maior consistencia transacional: a replicação *transacional* ou a *snapshot*?
3. Qual tem menor desfasamento: a replicação *merge* ou a *snapshot*?
4. O que distingue os modos *pull* vs *push subscription*?

Complementos de Bases de Dados – Replicação –

Engenharia Informática

2º Ano / 1º Semestre

Cláudio Miguel Sapateiro

claudio.sapateiro@estsetubal.ips.pt