Utilização da notação BPMN na refatoração de bancos de dados.

Veja como é possível utilizar a Business Process Modeling Notation (BPMN) no auxílio da refatoração de um banco de dados.

**Matheus Vinícius Mahl**

[matheusmahl@outlook.com](mailto:matheusmahl@outlook.com)

**Thyago César Souto**

[thyago\_tcs@yahoo.com.br](mailto:thyago_tcs@yahoo.com.br)

|  |
| --- |
| **De que se trata o artigo** |
| Apesar da refatoração de bancos de dados já não ser um assunto muito novo, poucos administradores ou mantenedores de bancos de dados o utilizam com regularidade. Esse processo acaba sendo pouco utilizado por diversos motivos, dentre eles, a dificuldade de se realizar operações de refatoração em um esquema que já está em funcionamento. Porém é possível utilizar ferramentas de modelagem para facilitar esse processo, principalmente na parte inicial da refatoração, onde é difícil identificar e tomar decisões sobre o que é preciso refatorar. Será abordado nesse artigo a utilização da modelagem de processos como um auxílio na refatoração, tornando o processo mais simples e de fácil entendimento. |
| **Em que situação o tema é útil** |
| Utilizando a modelagem do processo de refatoração descrita nesse artigo, a tarefa de refatoração do banco de dados tornar-se-á uma atividade mais simples e de melhor entendimento por parte do administrador do banco de dados e também dos desenvolvedores que farão a manutenção dos sistemas que utilizam o banco. Todo o processo de refatoração do banco ficará melhor documentado e melhor gerenciado, evitando falhas e prevenindo erros, permitindo que o processo seja feito mantendo a integridade dos dados. |

A refatoração de banco de dados por notação BPMN possibilita ao usuário criar um diagrama de processos de negócios orientado diretamente para uso humano. A função principal da modelagem é a automatização de processos e identificação de possíveis falhas existentes no processo, possibilitando ajustar o percurso em busca da otimização. Também permite que o processo se torne de fácil entendimento, consequentemente tornando a refatoração mais dinâmica e de maior compreensão. Utilizando essas teorias, esse artigo busca facilitar a refatoração do banco de dados.

Em primeiro momento, será conceituado um pouco sobre refatoração e sobre métodos formais, em específico sobre a notação BPMN. O foco do artigo não é se aprofundar em algum desses assuntos, mas utilizar ambos em conjunto para facilitar o processo de refatoração.

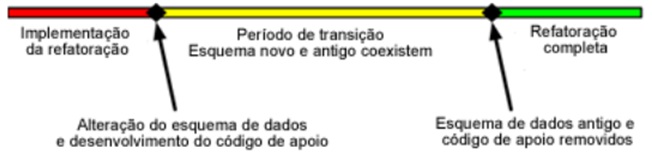
Refatoração

Refatoração de código é uma maneira disciplinada de reestruturação do código fonte em pequenos passos, sempre buscando evoluir e melhorar a qualidade de seu design. A refatoração tem como base a manutenção do código sem alterar o seu comportamento, ou seja, não adiciona e nem remove funcionalidades do seu projeto. Na mesma linha de pensamento, a refatoração de banco de dados é uma pequena mudança no esquema do banco que aprimora o seu design, enquanto mantém tanto o seu comportamento semântico como as informações contidas nele.

A refatoração de banco é conceituadamente mais difícil que a refatoração de código, já que é necessário manter tanto a semântica comportamental como a informacional. Adicionalmente a isso, existe a possibilidade de o banco de dados ser muito acoplado com outros programas externos, tais como frameworks, outras aplicações que o utilizam em paralelo, outros databases, etc.

Quando à necessidade de alteração em um banco é mencionada as demais aplicações vinculadas a esse banco sofrem pequenas alterações, entendendo que essas alterações simultâneas não serão implantas no mesmo momento, o ideal seria criar um período de transição para que a refatoração possua um prazo de inicio e fim e estejam alinhadas uniformemente à alteração e a versão original. Na **Figura 1** a situação mencionada anteriormente é relatada.

**Figura 1: - Ciclo de vida de uma Refatoração**



Fonte: Adaptado de Domingues (2014)

Para compreender e facilitar os ajustes necessários em diversos bancos de dados Ambler e Sadalage mencionam alguns tópicos e subitens, os quais seguidos rigorosamente possibilitam uma compreenção mais clara sobre o assunto e uma possível solução para o problema.

As praticas para refatoração estão divididas em quatro grupos: estrutural, qualidade de dados, integridade referencial e integridade arquitetural.

**I Estrutural**

A refatoração estrutural tem como objetivo principal a busca pelo melhor resultado no modelo do banco de dados. As principais causas para alterar a estrutura de um banco de dados são as mudanças e alterações nos requisitos do projeto, as alterações nos requisitos quando realizadas, geram mudanças por todo o projeto. A refatoração estrutural quando necessária, incluem alterações como:

* Remover tabelas, colunas ou visões.
* Dividir ou unir colunas e tabelas.
* Renomear colunas, tabelas ou visões.

**II Qualidade de Dados**

A refatoração de qualidade de dados possui o objetivo principal de diminuir as inconsistências presentes no banco de dados e por outro lado aumentar as validações necessárias presentes no banco de dados. Na refatoração de qualidade de dados encontramos ações referentes à:

* Padronizar colunas, aplicar formato comum e introduzir restrição à coluna desejada.
* Aplicar padrão e tipo ao código.
* Consolidar e implementar estratégia de chaves padronizadas.
* Remover as restrições de colunas, eliminar valor padrão ou restrição de not null.
* Possibilitar as colunas receberem o valor nulo.

**III Integridade Referencial**

A refatoração de integridade referencial possui como objetivo principal a consistência das referências com a sua tabela em questão. Alterações como remoção e inclusão de regras, armazenamentos históricos e também adições de lógicas da aplicação no banco de dados devem surtir o efeito correto no banco não comprometendo funções já adicionadas. A refatoração de integridade referencial possui ações referentes à:

* Inserção, remoção e alteração de chaves estrangeiras.
* Utilizar a exclusão em cascata adicionando e removendo funções.

**IV Integridade Arquitetural**

A refatoração Arquitetural tem como objetivo principal mover a parte de código dos sistemas anexados ao banco de dados para o próprio banco de dados, a parte lógica ficará dentro do próprio banco de dados. Como objetivo secundário são mencionados: a organização, a padronização e a melhora efetiva no desempenho do banco de dados em questão. A refatoração Arquitetural inclue ações como:

* Adicionar uma tabela de espelho, um uso de índices e também o uso de tabelas que possuam somente leitura.
* Adicionar métodos ao banco como *StoredProcedures*.

O objetivo principal de um processo é identificar uma série ou conjunto de atividades necessárias e que devem ser realizadas para alcançar o objetivo almejado. Definir a ordem em que as tarefas levantadas anteriormente sejam realizadas corresponde também a uma identificação e melhoramento da qualidade.

O conceito básico e principal de modelagem é criar uma sequência de processos ou um “modelo” representados através de diagramas operacionais e desenhos expressando o comportamento do banco de dados. A modelagem garante eficácia no projeto validando e testando suas reações sobre diversas condições com o intuito de que seu funcionamento saia como o desejado, esperando atender a todos os requisitos estabelecidos pelo projeto. Sua função atribuirá ainda mais valor quando houver a validação de desempenho e custo no ponto de vista de qualidade de produto.

Não existe um padrão para modelo de processos, o modelo em questão pode conter um ou mais diagramas como: informações referentes a objetos; informações referentes a relacionamentos entre os objetos; informações das atividades do processo com pessoas; informações das atividades do processo com sistemas de informação e também as funções automatizadas de processos.

Referente a técnicas de modelagem de processos, serão destacadas como as mais difundidas e conhecidas atualmente: Fluxogramas, BPMN (*Business Process Modeling Notation*), UML (*Unified Modeling Language*), IDEF0 e IDEF3 (*IntegratedDefinition*), e as Redes de Petri. A BPMN é a notação mais amplamente utilizada para modelar os processos.

Notação BPMN

No ano de 2005, a BPMNI (*Business Process Management Initiative*) deu inicio ao processo de desenvolvimento da notação gráfica originalmente conhecida e chamada por BPMN. Logo em seguida, a OMG continuou o processo de desenvolvimento da BPMN onde a mesma foi unida à notação XPDL considerando assim a notação BPMN como uma linguagem formal. Essa união contribuiu para o desenvolvimento e utilização de novas ferramentas, desde que as novas ferramentas utilizem os padrões publicados pela OMG.

*Business Process Model and Notation*, ou BPMN, é uma notação gráfica de modelagem de processos de negócios, cujo objetivo é prover uma gramática de símbolos para mapear os processos de negócios de um projeto ou organização. Essa notação é dita também como orientada para uso humano, referente a sua fácil compreensão e fácil identificação dos padrões de diagrama formado.

Com essas características, essa modelagem consegue dar suporte ao gerenciamento de processos de negócio, tornando-se capaz de representar semânticas de processos complexos e facilitando o entendimento do processo, tanto para o usuário técnico quanto para os usuários de negócio.

Destacando como ponto positivo da notação BPMN o fato de que os diagramas produzidos possam ser integrados a um ambiente operacional e automatizado. Outro ponto positivo é possibilitar que as linguagens XML desenhadas para execução de processos de negócios possam ser visualizadas e utilizadas dentro de uma notação orientada a processos.

A facilidade em modelar utilizando BPMN está atribuída à praticidade da notação em modelar através de processos no intuito de ser voltada diretamente a modelagem humana. A importância em modelar utilizando a notação BPMN é destacada em quatro itens:

* BPMN é um padrão internacional de modelador de processos aceito e utilizado pela comunidade.
* A BPMN será útil independentemente de qualquer metodologia de modelador de processos utilizada.
* Uma ponte padronizada será criada pela BPMN com o objetivo de diminuir a lacuna entre os processos de negócios e sua implementação.
* A BPMN de maneira geral proporciona uma modelagem unificada e padronizada do processo.

**Conceitos Básicos da Notação BPMN**

Na notação BPMN existem vários elementos para se construir uma modelagem de auto nível porem, os elementos mais utilizados para realizar esses processos de negócios são quatro: atividades (*tasks*), eventos (*events*), decisões (*gateways*) e os sequenciais de fluxo (*sequenceflows*), como mostra a **Figura 2**. Utilizando apenas esses métodos é possível construir modelos de processos bem elaborados e com as funcionalidades necessárias para representar graficamente.

**Figura 2: - Elementos básicos da notação BPMN**

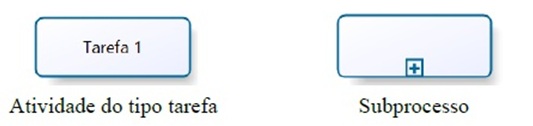


Fonte: Adaptado de Domingues (2014)

Como cada elemento possui suas funções e características, será destacada sua funcionalidade e exemplificar brevemente em que situação cada um pode ser usado.

* Atividade (*activity*): Atividade pode ser entendida por um termo do trabalho realizado em um processo. Uma atividade pode possuir outras atividades sendo assim chamada de Atividade Atômica, se possuir um processo ou um subprocesso dentro da atividade será denominada Atividade Composta. Na **Figura 3** temos um exemplo perfeito de Atividade Composta.

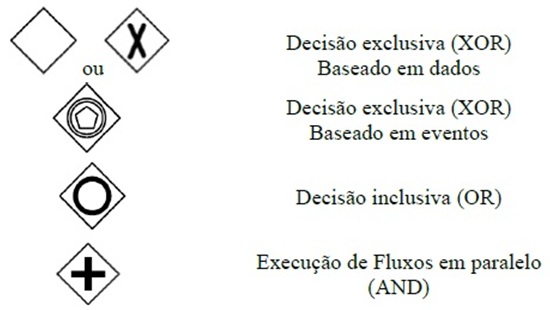
**Figura 3: - Tipos de Atividade da notação BPMN**



Fonte: Adaptado de Domingues (2014)

* Evento (*event*): Uma definição básica de evento é uma ação que acontece durante o desenvolver de um processo e por consequência altera o seu fluxo. O evento em si é dividido em três tipos diferentes de eventos: o evento de inicio; o evento intermediário; e o evento de fim.
* Decisão (*gateway*): A decisão é um elemento que possui a habilidade de mudar o caminho do processo, sua principal característica é alterar estados do processo. Sendo um elemento de controle de sequencia e interação dentro de um processo o gateway ou decisão, pode ser interpretado como uma pergunta que é feita dentro do processo, dependendo de sua resposta o processo pode ir para outro estado. Na **Figura 4** serão apresentadas as várias formas de decisão.

**Figura 4: - Tipos de *Gateways* da notação BPMN**



Fonte: Adaptado de Domingues (2014)

* Conector de Sequencia de fluxo (*Sequenceflow*): Um fluxo de sequencia possui a característica e habilidade de coordenar em que ordem as atividades do processo serão executadas. O conector indica que há uma ligação gerada através de dependência e prioridade entre as tarefas envolvidas, determinando seu inicio e fim. A sequencia de fluxo é responsável por ligar os demais elementos do processo, ela conecta atividades, decisões e eventos uns aos outros. As atividades possuem a característica de ter somente uma única entrada e uma única saída.

**Diagramas de Processo de negócio (BPD)**

O BPD (*Business Process Diagram*) possui a finalidade de auxiliar na comunicação entre uma ampla variedade de informações e uma ampla variedade de publico. O BPMN possibilita a utilização de muitos tipos de modelagem e a criação de um processo de negócio com começo e fim respeitando todos os requisitos. Os elementos estruturais do BPMN permitem ao utilizador uma identificação exata e clara do processo por completo passando de diagrama a diagrama e identificar facilmente o seu fluxo.

Na notação BPD (*Business Process Diagram*) existem três tipos de processo de negocio como: *private; abstract e colaboration.*

**Private (*Internal Bussines Process*)**

Mais conhecidos como processos de negócios privados, são utilizados quando não há interesse na interação do processo com os demais processos que ele possa interagir. O objetivo principal é o real fluxo do processo.

**Abstract (*Public Process*)**

Conhecidos como processos abstratos, representam uma interação entre um processo de negocio participante e um processo de negocio privativo. A real preocupação do processo abstrato é como ele colabora com os outros processos ligados a ele, o fluxo é relevante no *Abstract Process.*

**Colaboration (*Global Process*)**

O processo colaborativo representa a interação entre duas ou mais entidades de negocio, mencionadas como sequencias de atividades essas interações representam o padrão de troca de mensagens entre todas as atividades envolvidas no processo.

Semelhante a idéia de comunicação o processo colaborativo pode ser entendido como a união de dois ou mais processos mantendo uma comunicação entre si.

[CHECKPOINT]

Modelagem de processos na refatoração do banco de dados

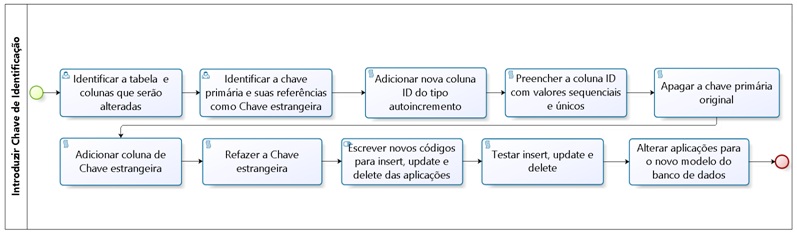
A complexidade das alterações feitas em uma refatoração de banco de dados pode ser diminuída com a ajuda da modelagem de processos. Utilizando a notação BPMN é possível realizar uma análise do processo de refatoração, separando, classificando e ordenando as alterações que serão feitas no banco de dados, tornando o processo mais organizado e de fácil compreensão.

De início, é necessário coletar as informações necessárias para realizar a modelagem. Portanto é preciso identificar os tipos de refatorações que deverão ser aplicadas ao banco, conforme a necessidade identificada em se realizar a refatoração. A seguir será descrito alguns casos de refatorações que podem ser realizadas no banco, as quais foram escolhidas por serem as mais comumente realizadas.

**Introduzir chave de identificação**

Em alguns momentos, é possível se deparar com tabelas que utilizam múltiplas chaves primárias em vez de uma única chave identificadora, normalmente auto incrementável. Portanto, caso seja vista a necessidade de se realizar a refatoração dessa tabela para remover as chaves primárias compostas, pode ser utilizado uma sequência de passos, como mostrado na **Figura 5**.

**Figura 5 – Sequência de passos para introdução da chave de identificação**



Fonte: Adaptado de Domingues (2014)

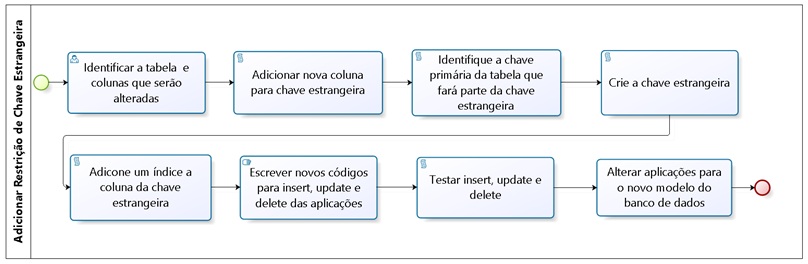
É possível perceber que a utilização da notação BPMN para a modelagem do processo o torna mais compreensível. Seguindo a sequência de passos corretamente, é possível ter uma segurança maior ao fazer as alterações, sendo que é possível desfazer as alterações caso ocorra algum problema.

Também é possível notar que realizando essa operação, também é necessário realizar a refatoração da chave estrangeira nas tabelas relacionadas a ela.

**Refatoração de chave estrangeira**

Outro tipo de refatoração é a adição ou alteração de chaves estrangeiras (ou Foreign Key). Esse tipo de refatoração é muito útil para realizar o relacionamento entre as tabelas do banco. Caso seja identificada a necessidade da se impor uma restrição de chave estrangeira no banco, é possível seguir a modelagem descrita na **Figura 6.**

**Figura 6 – Sequência de passos para a adição de restrição de chave estrangeira**



Fonte: Adaptado de Domingues (2014)

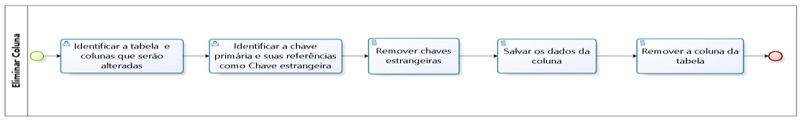
**Testes**

Lembre-se sempre de realizar testes, antes e depois das alterações. Isso garantirá que as refatorações não tenham alterado o comportamento ou a semântica do banco de dados. Durante a refatoração em ambiente de homologação, caso ocorra um erro ainda é possível voltar atrás e recuperar os dados.

**Refatoração de eliminação de coluna**

Essa refatoração é importante ser realizada quando a coluna de uma tabela não é mais utilizada, principalmente por consequência de outras refatorações. Quando foi identificada a necessidade de remoção de uma coluna, essa deve ser realizada para evitar que ela seja usada por engano. A eliminação de uma coluna é uma refatoração fácil de ser realizada, porém ainda é necessário tomar cuidados na realização dessa tarefa, para evitar que dados sejam perdidos ou ocorram erros nas restrições, como chaves estrangeiras ligadas à coluna. Na **Figura 7** é possível ver quais passos devem ser seguidos para realizar essa refatoração.

**Figura 7 – Passos para a Refatoração de eliminação de colunas.**



Fonte: Adaptado de Domingues (2014)

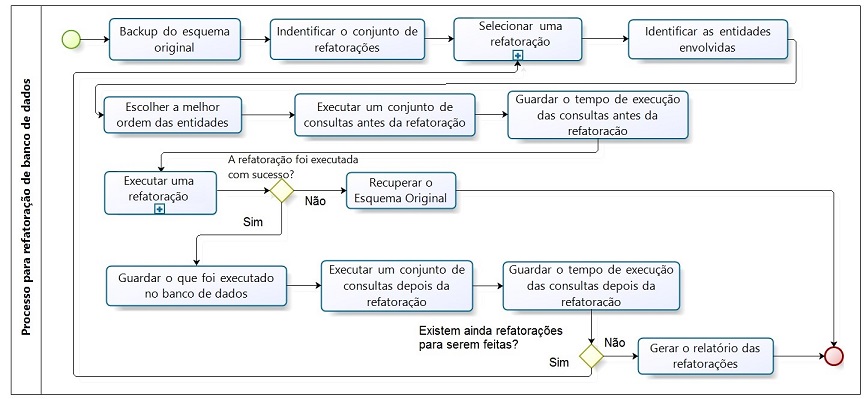
Além dos tipos de refatoração apresentados, é possível criar modelagens dos outros tipos de refatoração que podem ser necessários serem realizadas no banco de dados. Como essas refatorações podem ser aplicadas diversas vezes em vários tipos de bancos, é possível criar uma modelagem padrão e utilizá-la nas diversas refatorações que um banco pode sofrer.

É possível adicionar passos conforme a necessidade, como a de realizar testes em diferentes etapas da refatoração, colocando isso em um processo, o qual pode facilitar o entendimento do processo de refatoração e auxilia na realização correta das alterações, evitando problemas.

Grandes alterações no banco de dados

As modelagens apresentadas anteriormente representam pequenas alterações de refatoração em um banco de dados. Normalmente é necessário realizar diversas alterações em um mesmo esquema para que a refatoração seja feita com sucesso. Esse processo é conhecido como uma grande alteração no banco de dados e pode ser muito complexa de ser realizada.

Para garantir que todas as etapas da grande refatoração sejam realizadas, é preciso analisar e decidir quais alterações devem ser realizadas antes das outras. Para facilitar o entendimento e execução da sequência corretamente é possível modelar esse processo, assim como foi feito com as pequenas refatorações. É possível ver na **Figura 8** um exemplo de como realizar a modelagem da sequência de tarefas que devem ser realizadas durante a refatoração.

Figura 8 – Processo para refatoração de banco de dados.

Nesse modelo de processo podemos notar que a refatoração segue uma sequência lógica e permite que diversas refatorações possam ser feitas, desde que sejam identificadas e ordenadas corretamente.

Seguindo os passos, a primeira coisa a se fazer é o backup da base de dados original. Mesmo que seja em ambiente de homologação, é importante que essa tarefa seja realizada para garantir a recuperação do esquema alterado.

O segundo passo é a identificação do conjunto de refatorações. Essa tarefa pode ser realizada a partir de um conjunto de informações sobre problemas ou sugestões identificadas em ambiente de produção. Utilizando essas informações é possível montar uma lista de refatorações a serem realizadas no banco. É preferível também que essa tarefa seja realizada por um administrador de banco de dados que tenha amplo conhecimento sobre o esquema a ser alterado. Assim que for identificado o que deve ser feito e montado um processo para a execução, as tarefas seguintes podem ser executadas até mesmo por um administrador que não possua muita experiência para verificar os resultados.

Os processos seguintes para a refatoração buscam guiar o administrador durante a refatoração mostrando os passos a serem seguidos. Realizando a execução dos processos em ambiente de desenvolvimento, é possível ganhar experiência sobre a ordem de execução das refatorações e sobre como melhorar o processo de refatoração. Os exemplos apresentados nesse artigo permitem que seja realizada a refatoração de diversos tipos de bancos de dados, além da expansão dos processos para que melhor se aplique na realidade do banco que precisa ser refatorado.

Conclusão

A notação BPMN mostrou-se apta a auxiliar no processo de refatoração de banco de dados, todas as refatorações descritas e exemplificadas anteriormente possuem passos definidos e concretos que podem ser facilmente representados pela notação BPMN. Utilizando a notação existe a possibilidade de modelar as restrições necessárias, estabelecer a ordem em que as tarefas serão realizadas e garantir a rastreabilidade onde posteriormente, as alterações realizadas poderão ser revertidas voltando ao estado inicial garantindo a integridade do banco.

Desta maneira entendemos que os processos e exemplos apresentados nesse trabalho podem ser utilizados em bancos de dados onde o DBA identifique que exista a real necessidade de melhor o banco em questão. A utilização da notação BPMN além de agilizar o processo de refatoração pelo DBA também, contribui para identificação de erros, garantem a integridade e garantem a consistência do banco de dados.

**Links**

DOMINGUES, Márcia Beatriz Pereira; ALMEIDA JUNIOR, Jorge Rady de; PALETTA, Francisco Carlos. **MODELAGEM DE PROCESSOS DE REFATORAÇÃO DE BANCO DE DADOS UTILIZANDO BPMN.**São Paulo: Usp, 2014. Disponível em: <http://www.contecsi.fea.usp.br/envio/index.php/contecsi/12CONTECSI/paper/viewFile/3007/2342>. Acesso em: 19 jun. 2016.