

Requisitos para Ferramentas de Migração de Dados

Pedro de Alcântara dos Santos Neto¹, Josino Rodrigues Neto², Francisco das Chagas Ribeiro Júnior³, Pedro Almir Oliveira¹

¹DC - UFPI – Teresina, PI – Brasil

²CIN – UFPE – Recife, PE – Brasil

³ Infoway Tecnologia e Gestão em Saúde Ltda, Teresina, PI – Brasil

pasn@ufpi.edu.br, jrn4@cin.ufpe.br, junior@infoway-pi.com.br,
petrus.cc@gmail.com

Abstract. *This work presents a requirements catalog for data migration tools. It captures the needs, desire and expectations related to users of migration tools, based on our experience dealing with migration projects. The catalog can be used to facilitate the development of new tools. Besides, the catalog can be used as basis for comparison among tools, facilitating the identification of suitable tools for specific projects. In this work is still presented a comparison of six tools using the catalog.*

Resumo. *Este trabalho apresenta um catálogo de requisitos associados a ferramentas de migração de dados. Ele captura as necessidades, desejos e expectativas relacionadas aos usuários de ferramentas de migração, baseado em nossa experiência em projetos envolvendo migrações, e pode ser utilizado para facilitar o trabalho de desenvolvimento de novas ferramentas que auxiliem nesse processo. Além disso, o catálogo pode ser utilizado como base para a comparação de ferramentas existentes, facilitando a identificação de produtos mais apropriados a certos projetos. Para tanto, apresenta-se uma comparação de seis ferramentas disponíveis no mercado, utilizando como base o catálogo desenvolvido.*

1. Introdução

A quantidade crescente de diferentes tecnologias para gerenciamento de dados faz com que a melhor opção para armazenamento em uma determinada época esteja obsoleta noutra. Em casos como esse, a migração de dados para um sistema mais moderno é a decisão tomada por muitos gerentes de projeto. A finalidade de um processo de migração, seja por motivo de troca de sistema ou por troca de Sistema Gerenciador de Banco de Dados (SGDB), é atender às necessidades atuais e futuras da empresa (Stair e Reynolds, 2002).

Segundo pesquisa da Bloor Research (Howard e Potter, 2007) com a participação de 700 empresas da Forbes Global 2000, 84% dos projetos de migração falham no aspecto orçamento (ficam mais caros do que o planejado) ou são entregues muito além do prazo determinado.

Partindo do princípio de que projetos de migração são elementos-chave em projetos de aplicações maiores e mais abrangentes, é essencial a adoção de disciplinas, metodologias e ferramentas adequadas para que eles não fracassem. A conveniência do processo de migração de dados justifica a criação de ferramentas e técnicas que auxiliem

a realização dessas tarefas em um menor tempo e com menos esforço por parte das equipes de desenvolvimento.

Acreditamos que a grande contribuição para qualquer área seja a especificação de um catálogo de requisitos que capte os desejos, necessidades e expectativas de usuários em relação a tal área ou ferramenta, conforme já feito por outros autores (Santos-Neto et al, 2005). Os catálogos de requisitos auxiliam profissionais a selecionar métodos e ferramentas, de conforme suas expectativas.

Neste trabalho apresentamos os requisitos que devem ser levados em consideração para ferramentas de migração de dados. Os requisitos listados foram obtidos por meio da experiência adquirida a partir da execução de projetos de migração de dados, reforçados pela criação de uma ferramenta de migração e ainda fundamentados em uma vasta pesquisa bibliográfica sobre ferramentas e métodos existentes. Assim, o trabalho apresenta como contribuições principais: i) o levantamento de uma lista de requisitos associados às ferramentas de migração, facilitando o trabalho de desenvolvimento de novas ferramentas que auxiliem nesse processo; ii) a listagem de alguns problemas que podem ser encontrados durante projetos de migração; iii) a criação de um mecanismo simples para comparação de ferramentas existentes, promovendo a identificação de produtos mais apropriados a certos projetos.

Este estudo está organizado da seguinte maneira: na Seção 2 são discutidos os trabalhos relacionados que de alguma forma contribuíram para o desenvolvimento deste artigo; na Seção 3 é apresentada a lista de requisitos identificados; na Seção 4 são analisadas algumas ferramentas existentes, com uma breve comparação entre elas, utilizando como base os requisitos identificados; e, finalizando, na Seção 5 são expostas as conclusões do trabalho.

2. Trabalhos Relacionados

As ferramentas de migração de dados disponíveis são várias. A grande maioria delas está associada a alguma tecnologia específica e serve apenas para ajudar na carga de dados. Conforme veremos a seguir, isso está associado a apenas um dos requisitos existentes para ferramentas dessa natureza.

Existem vários trabalhos que abordam metodologias para migração de dados e uma série de outros trabalhos tratando especificamente de testes e qualidade de dados após a migração (Ramalho et al 2012, Matthes et al 2011, Haller et al 2008, Fontanette et al, 2004). Ramalho et al (Ramalho et al 2012) fazem um apanhado geral das metodologias propostas para migração de dados e apresentam uma classificação de abordagens para migração de bancos de dados e uma metodologia unificada originada a partir das diversas abordagens de migração referenciadas pelos autores. A Figura 1 apresenta a classificação proposta pelos autores. Matthes et al (Matthes et al, 2011) apresentam um outro modelo para migração de dados e técnicas de testes e garantia da qualidade dos dados durante e após a execução da migração. O Modelo de processo proposto e as considerações a respeito de testes foram construídos com base na larga experiência dos autores com a execução de migrações de dados. Este trabalho, entretanto, se diferencia dos supracitados por focar no processo de transferência de dados entre sistemas de armazenamentos diferentes.

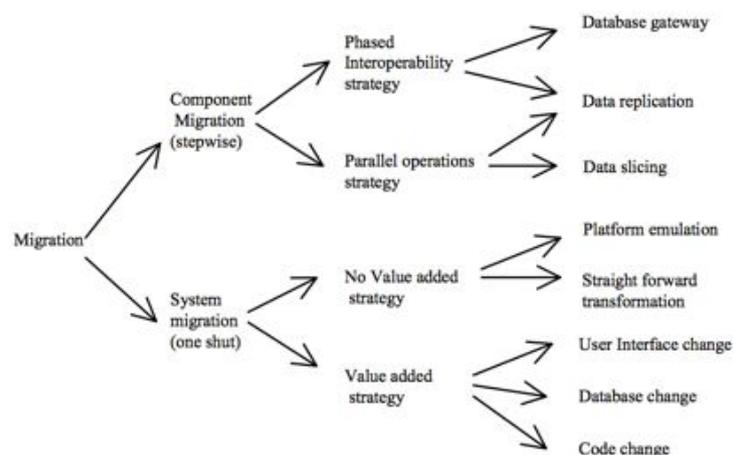


Figura 1: Classificação de abordagens para migração de repositórios (Ramalho 2012).

Não encontramos trabalhos com o intuito de avaliar, com base em critérios bem definidos, ferramentas de migração. Da mesma forma, não nos deparamos com trabalhos visando criar um catálogo de requisitos associados a ferramentas de migração de dados. No entanto, encontramos trabalhos relacionados à criação de catálogo de requisitos para o teste baseado em modelos (Santos-Neto et al, 2005), assim como para ferramentas de gerenciamento de requisitos (Hoffmann et al, 2004).

Por conta da ausência de trabalhos nessa área, acreditamos que as contribuições listadas aqui, embora em caráter preliminar, certamente ajudarão os usuários e desenvolvedores de ferramentas de migração de dados.

3. Requisitos para ferramenta de migração de dados

Esta seção apresenta os requisitos para ferramentas de migração de dados. A estrutura utilizada para organizar a seção é a seguinte: o nome da subseção indica o requisito e o texto da subseção traz a discussão dos diversos aspectos relacionados a ele, evidenciando problemas e soluções associadas ao que está sendo discutido.

O principal requisito associado à migração de dados é a possibilidade de realizar a migração, que se refere ao processo de transferir dados entre sistemas de armazenamentos diferentes. No entanto, focamos nos requisitos adicionais, que estão ligados a essa necessidade básica, mas vão além disso.

Ressaltamos que a migração de dados normalmente é uma tarefa dispendiosa e necessita de uma equipe para sua execução. Por conta disso, é fundamental que ela seja tratada como um projeto, devendo ter escopo, cronograma, acompanhamento e tudo o que um projeto necessita. Porém, essas questões não são tratadas pelas ferramentas de migração, uma vez que elas são focadas na transferência dos dados. Deste modo, não existe nenhum requisito ligado a essa questão, visto que nosso foco está nas ferramentas de migração.

3.1. Migração envolvendo estruturas diferentes

O uso de aplicativos de escritório é algo comum atualmente. Diversas pessoas utilizam editores de texto e planilhas eletrônicas como ferramentas de trabalho, registrando informações úteis para seus processos de negócio. Por conta disso, é importante que as

ferramentas de migração possuam algum suporte para a transferência de dados entre estruturas diferenciadas.

Conforme já comentado, o principal requisito relacionado a uma ferramenta de migração é sua habilidade de transferir dados entre sistemas de armazenamentos diferentes. Isso pode envolver o uso de estruturas alternativas, como migrar dados de planilhas eletrônicas, arquivos CSV (*comma separated values*) e arquivos textos para bancos de dados.

A existência de um suporte para a migração envolvendo estruturas diferentes permite mais agilidade ao processo, com maior controle da qualidade de sua execução, além de maior abrangência no uso de informações. Com isso, os sistemas de informação tendem a ficar mais completos, pois podem conter o agrupamento de todos os dados relacionados à sua área de atuação, independente da origem das fontes para transferências desses dados.

3.2. Migração paralela ou distribuída

Em boa parte dos projetos de migração existe uma grande quantidade de dados para serem migrados. Nesses casos, a busca de alternativas que minimizem o tempo de execução das migrações é algo vital para o sucesso do projeto. Uma solução possível para esse problema é o uso de programação paralela ou distribuída durante a execução da migração, permitindo assim que tarefas independentes possam ser executadas em um tempo menor, utilizando recursos computacionais diferentes.

O uso de paralelismo em projeto de migração de dados não é algo novo. Lübeck em (Lubeck et al, 2003) apresenta sua experiência relacionada à migração de dados no CERN (*European Organization for Nuclear Research*). O desafio nesse projeto era realizar a migração de 300TB de dados. Em sua abordagem, Lübeck utilizou um *framework* para distribuir as tarefas de migração entre vários nós que executariam paralelamente e que eram monitorados e controlados por um aplicativo que executa em um nó separado.

Sempre que admissível essa pode ser uma alternativa viável para a execução da migração em um espaço de tempo reduzido, portanto, é importante que haja suporte das ferramentas de migração para migração paralela ou distribuída.

3.3. Migração incremental

Quando o sistema que se pretende migrar é de grande porte e possui muitas funcionalidades vitais para a instituição que o utiliza, nem sempre é possível realizar o processo de migração por inteiro. Nesses casos, uma abordagem viável é a migração gradativa da aplicação legada. Fontanette et al (Fontanette et al, 2004) realizaram uma pesquisa em que foi apresentada a viabilidade da migração gradativa de aplicações legadas.

A migração de uma aplicação legada pode ser feita de forma direta ou gradativa. Na abordagem direta, a aplicação passa pelo processo de transferência de dados e somente é disponibilizada para a empresa ao final desse processo. Na abordagem gradativa a aplicação legada é convertida por módulos. Durante o período em que a aplicação legada passa pelo processo de migração, a empresa continua a utilizar a

aplicação legada e, de acordo com a disponibilização dos módulos convertidos, disponibiliza-se também a aplicação nova, com os dados migrados.

A migração gradativa proporciona para instituição que a adota maior flexibilidade, resultados em curto prazo, mitigação dos riscos, planejamento e maior controle do processo de migração. Durante um processo de migração é imprescindível o uso de uma ferramenta que automatize a migração e minimize a probabilidade de erros humanos, pois a possibilidade de se fazer migrações incrementais pode ser fundamental em alguns projetos.

3.4. Reengenharia de dados

O processo de análise, reorganização de estruturas e dos valores dos dados, para torná-los mais compreensíveis, é chamado de reengenharia de dados (Sommerville, 2003). A reengenharia de dados é um processo utilizado para padronizar dados de uma base para um formato comum, corrigindo problemas de qualidade, removendo duplicação de informação, construindo novos relacionamentos entre tabelas ou enriquecendo a base de dados com informações suplementares.

Aebi e Largo (Aebi e Largo, 1994) relatam uma série de problemas associados à reengenharia de dados, dividindo-os em duas categorias: problemas de definição de dados e problemas de valores de dados. Como principais problemas de definição de dados, destacamos:

- Mapeamento de esquema - atributos e tabelas do esquema antigo precisam ser mapeados para o novo esquema de acordo com os requisitos da nova modelagem.
- Consistência de restrições (*constraints*) - a nova base de dados pode ter restrições novas, alteradas ou eliminadas, em relação à base de dados legada.
- Tabelas e relacionamentos - tabelas novas podem ser introduzidas ou relacionadas com as tabelas já existentes.
- Heterogeneidade de chaves-primárias - tabelas são identificadas por diferentes construções de chaves.

Os problemas de valores de dados estão diretamente ligados ao conteúdo existente nos mecanismos de armazenamento. Dentre esses problemas destacamos:

- Identificação de objetos - algumas entidades podem aparecer duplicadas no mesmo contexto.
- Conflito de limites - o valor de um atributo pode não ser compatível com o limite definido. Por exemplo, poderia haver um valor negativo em um campo que deveria conter somente valores positivos.
- Diferenças de unidade, escala e granularidade - As unidades de medida do sistema novo podem diferir do sistema antigo.
- Codificação - valores de dados frequentemente podem possuir definição de codificação diferente.

Uma boa ferramenta de migração deve possuir formas de resolver esses tipos de problemas. Para alguns projetos de migração, uma ferramenta que auxilie na resolução desses problemas pode ser indispensável.

3.5. Migração envolvendo paradigmas diferentes

Há certo tempo, os bancos de dados relacionais eram o padrão de fato no desenvolvimento de aplicações corporativas e aplicações *web*. Com o advento da computação em nuvem começaram a surgir novos paradigmas de armazenamento de dados. O surgimento desses novos paradigmas tem como objetivo viabilizar o melhor uso dos recursos existentes na Internet (*cloud*), facilitando a escalabilidade para níveis considerados muito altos e difíceis de serem tratados diretamente pelos bancos de dados relacionais.

Existem diversos Sistemas de Gerenciamento de Banco de Dados (SGBDs) em nuvem, cada um com características e propósitos específicos. O fato de serem sistemas de gerenciamento de dados relativamente novos faz com que existam poucas ferramentas que forneçam suporte a migração de dados para esses paradigmas. No entanto, devido ao crescimento do seu uso pela comunidade em geral, é importante que haja suporte das ferramentas de migração para transferências de dados envolvendo paradigmas diferentes.

3.6. Migração entre bases de dados com modelagens diferentes

Em projetos de migração de dados podemos ter 2 cenários. No primeiro cenário o objetivo é apenas mudar de tecnologia de armazenamento de dados. Por exemplo, mudar do SGBD Mysql para o Postgresql. Nesse caso, não há mudanças no esquema do banco de dados que está sendo migrado. Os dados são apenas transportados do banco de dados origem para o banco de dados destino. No segundo cenário o objetivo é migrar os dados entre fontes de dados com paradigmas e/ou modelagens diferentes. Nesse caso é necessário fazer um mapeamento das bases de dados para que se saiba onde cada valor lido na fonte de dados origem será armazenado na fonte de dados destino.

Um exemplo prático desse fato é quando possuímos, por exemplo, os dados de pessoa física (nome, data de nascimento, identidade, endereço, etc.) armazenados em uma única tabela na fonte de dados origem e esses mesmos dados são armazenados em duas tabelas na base de dados destino, uma tabela para os dados de endereço e uma tabela com os outros dados da pessoa.

Para migrações entre fontes de dados com paradigma e/ou modelagem diferente o mapeamento é um dos pontos mais importantes. Uma ferramenta que pretenda atender a esse tipo específico de migração deve prover um mecanismo de mapeamento das fontes de dados que seja eficiente e flexível.

3.7. Testabilidade

Após a migração dos dados, é necessário certificar-se de que todos os atributos de todos os registros selecionados para serem migrados foram transferidos conforme o planejado. A forma mais comum de atingir tal objetivo é executar testes nos dados migrados.

Klauss Haller (Haller, 2008) expõe três estratégias para validação dos dados após a migração: i) *Top-down*: Consiste na construção de estatísticas com base na

contagem dos registros, após a migração, com o intuito de garantir que a quantidade de registros migrados está de acordo com o planejado; ii) *Botton-up equivalence*: Caracteriza-se pela verificação, com base na chave primária dos registros, da presença de todos os objetos escolhidos para serem migrados na base destino; iii) *Botton-up fingerprint*: Considera atributos chaves para a construção de uma identidade - uma espécie de código *hash* - para fazer a verificação dos dados migrados com base nesta informação. Assim, verifica não somente a presença dos registros, mas a migração correta de todos os seus atributos.

Jutta Horstmann (Horstmann, 2005) aborda a questão de uma forma mais geral, ressaltando que é necessário responder aos seguintes questionamentos:

1. Quantos registros esperamos que sejam criados?
2. Foi criada a quantidade certa de registros? Em caso negativo, por quê?
3. Os dados foram carregados nos campos corretos?
4. Os dados foram formatados corretamente?
5. A fonte de dados a ser migrada possui registros com atributos com valor nulo? Em caso positivo, eles foram migrados com sucesso?
6. Os campos numéricos foram migrados com a precisão correta?
7. As restrições de chave causaram algum erro? Em caso positivo, é possível determinar quais valores causaram tais erros.

Considerando os custos advindos de erros durante o processo de migração, uma ferramenta de migração deve viabilizar o teste da migração, uma vez que tal prática incorrerá em economia de tempo e dinheiro.

3.8. Boa usabilidade

Quando os autores deste trabalho realizaram o desenvolvimento e avaliação da ferramenta de migração de dados JExodus (Rodrigues-Neto e Passos, 2009), observaram que a usabilidade de uma ferramenta de migração poderia influenciar bastante a produtividade da equipe de desenvolvimento responsável pela migração de dados. Foi verificado que existe a necessidade de representações gráficas para que o gestor ou desenvolvedor do projeto possa ter uma visão geral de toda a massa de dados que está sendo migrada.

Nos momentos em que constatamos a necessidade de codificação, observamos que uma boa ferramenta integrada à ferramenta de migração aumentaria a produtividade dos desenvolvedores da equipe de migração.

O estudo sobre a usabilidade de ferramentas de migração de dados e sua influência na produtividade dos desenvolvedores poderá ser realizado como trabalhos futuros. Dessa forma poderemos obter dados estatísticos que possam comprovar com alto grau de precisão quais pontos relacionados à interface podem ter prevalência sobre a produtividade dos desenvolvedores.

4. Avaliação das Ferramentas

Com base nos requisitos levantados no presente trabalho, realizou-se uma avaliação de seis ferramentas de migração de dados. Escolheu-se como ferramentas uma open source, liberada sob a licença GPL, algumas comerciais e a ferramenta utilizada pelos autores em experiências de migração de dados. Uma descrição breve das ferramentas é dada a seguir:

- MySQL Migration Toolkit (<http://dev.mysql.com/doc/migration-toolkit/en/>): Ferramenta Open Source, integrante do pacote MySQL GUI tools, distribuída sob a licença GPL. Oferece possibilidade de migração de dados a partir de diversos SGBDs para MySQL.
- JExodus (Rodrigues-Neto e Passos, 2009): Ferramenta desenvolvida para realizar a migração de bases de dados tendo como fonte de dados um SGBD, documentos de texto ou objetos Java, de forma que o desenvolvedor efetue a migração de dados sem se preocupar com detalhes de baixo nível como conexão com bancos de dados, comandos SQL ou implementar código para ler dados de diversas fontes diferentes.
- SwisSQL (<http://www.swissql.com/>): Ferramenta comercial provida pela empresa Zoho. A ferramenta é capaz de efetuar migração a partir dos mais diferentes SGBDs e alguns tipo de arquivos (cvs, txt) para os mais diversos SGBDs.
- DBConvert (<http://dbconvert.com>). Ferramenta da empresa DMSoft Technologies, com a proposta de ser fortemente centradas sobre a migração de dados entre diferentes plataformas.
- DTM Migration Tool (<http://www.sqledit.com/mk/>). Ferramenta da empresa SQL Edit. Tal ferramenta se propõe a fazer a migração de dados entre diferentes bancos de dados de forma simples.
- Data Loader (<http://www.dblog.com/data-migration-tool.htm>). Empresa: Interface Computers. Uma ferramenta simples e poderosa para importar e exportar dados entre muitos formatos comuns.

4.1. Análise das ferramentas

As avaliações foram feitas com base na utilização das ferramentas em projetos experimentais de migração de dados, assim como uma leitura da documentação fornecida pelos seus mantenedores.

Os projetos experimentais para avaliação das ferramentas foram os seguintes:

- Controle de Curso com Alunos e Disciplinas. BD Original: MySQL. Paradigma do BD original: Relacional. Número de tabelas: 7. Número médio de registros por tabelas: 50. Ações a serem realizadas na migração: i) Na tabela aluno será retirada a colunas bairro, CEP, número e rua que irão compor outra tabela chamada endereço. ii) A tabela matrícula será apagada e a coluna matrícula será inserida na tabela aluno. iii) A dependência entre os dados pode tornar o processo de migração lento. Bancos de destino: Oracle, PostgreSQL, Firebird, MS Access, MS Excel. A operação será executada do banco de origem para os bancos de destino e vice-versa.
- Sistema de Biblioteca. BD Original: MySQL. Paradigma BD Original: Relacional. Número de tabelas: 4. Número médio de registros por tabelas: 50000. Ações para a migração: i) Transposição dos dados, sem mudanças nos formatos envolvidos. Bancos de destino: Oracle, PostgreSQL, Firebird, MS Access, MS Excel e CSV. A operação de será executada do banco de origem para os bancos de destino e vice-versa.

4.1.1. Migração envolvendo estruturas diferentes

A ferramenta MySQL Migration Toolkit não possui suporte ao uso de estruturas diferentes para transferência de dados. As ferramentas SwisSQL e JExodus permitem a transferência de dados envolvendo arquivos CSV. A SwisSQL, DbConvert, DTM e Data Loader possuem suporte à migração de dados a partir de planilhas eletrônicas, além de arquivos textos, CSV e outros formatos.

4.1.2. Migração paralela ou distribuída

A ferramenta DB Convert possui a capacidade de criar mais de um Thread para minimizar o tempo requerido no processo de migração. Nenhuma das outras ferramentas selecionadas para a análise dá suporte à migração paralela.

4.1.3. Migração incremental

O MySQL Migration Toolkit e o SwisSQL permitem a migração de apenas um conjunto de registros, além de armazenar os dados de um projeto já iniciado para posterior conclusão ou reuso.

O JExodus, a Data Loader e o DB Convert, possibilitam que um determinado grupo de objetos seja migrado por vez. O DB convert ainda ajuda o usuário, informando quais tabelas devem ser migradas juntas e quais não possuem dependências para serem migradas. Já o JExodus e a Data Loader deixam a cargo do usuário escolher como deseja migrar cada conjunto de registro e não fornece nenhuma ajuda a essa tarefa

A DTM Migration Tool não possibilita a migração incremental. Não é possível a criação de módulos para a migração.

4.1.4. Reengenharia de dados

O MySQL Migration Toolkit não possui suporte a tal tarefa durante a migração.

O SwisSQL permite reengenharia de dados por meio de funções escritas em SQL. A partir de tal recurso é possível gerar novas colunas com valores calculados com base em valores de outras colunas, utilizando rotinas SQL.

O JExodus possui suporte para a realização de uma reengenharia de dados a partir da criação de classes Java que especifiquem a rotina de transformação necessária.

A DB Convert possui ferramentas particulares para cada base de dados a ser migrada. Com isso a ferramenta evita inconsistência de valores padrões, problemas com a codificação, diferenças de unidades, escala e granularidade, conflito de limites e identificação de objetos redundantes.

A DTM e Data Loader não possuem recursos para ajudar a reengenharia de dados.

4.1.5. Migração envolvendo paradigmas diferentes

Nenhuma das ferramentas provê suporte para migração para bancos de dados com paradigmas diferentes, sejam os SGBDs destino orientados a objetos, NoSQL, etc.

4.1.6. Migração entre bases de dados com modelagens diferentes

O MySQL Migration Toolkit não possui suporte para tal tarefa durante a migração, uma vez que ele não permite a alteração do esquema dos dados em questão.

O SwisSQL permite a alteração do esquema do banco de dados a partir de recursos de sua interface gráfica, como também via manipulação, utilizando instruções SQL. O JExodus também permite que o esquema de dados destino seja alterado, porém, isso se dá via arquivos de configuração que detalham as modelagens – antiga e nova - a serem consideradas.

A DB Convert, DTM e Data Loader não possuem suporte para base de dados com modelagens diferentes.

4.1.7. Testabilidade

O MySQL Migration Toolkit apresenta de forma simplória algumas informações sobre a migração efetuada. A partir do relatório de dados fornecido pela ferramenta após o processo, verificou-se quantos objetos foram migrados em cada tabela. Porém, não informa quais registros não foram migrados e não permite a verificação da correção da migração de cada atributo.

O SwisSQL, DTM e Data Loader criam, ao fim da migração, um relatório no qual consta a quantidade total de registros, a lista de objetos migrados e a lista de objetos não migrados. Porém, assim como o MySQL Migration Toolkit, não permite a verificação de cada atributo individualmente, para verificar questões como: precisão, formato, etc.

O JExodus não fornece relatórios e nem possui suporte à verificação da precisão e formato dos dados após a migração. A DB Convert possui uma verificação de possíveis problemas a serem enfrentados na migração, mas não permite a execução de teste sobre os dados que estão sendo migrados.

4.1.8 Boa usabilidade

A MySQL Migration Toolkit, a Data Loader e a SwisSQL possuem interface gráfica para guiar seu uso, favorecendo a produtividade dos seus usuários. Outro fator relevante para a usabilidade da ferramenta SwisSQL é o fato dela tratar a migração como um fluxo com uma série de passos.

A ferramenta JExodus não possui uma interface gráfica associada. Foi constatado pelos autores deste trabalho que o uso dessa ferramenta pode ser mais oneroso que as demais, por conta do seu formato intensamente baseado em arquivos XML.

A DB Convert se preocupou em tornar fácil o processo de migração, o qual é feito em etapas bem assistidas pela ferramenta. O próprio tempo médio de execução das migrações (entre 8 e 10 minutos) mostra que a ferramenta é bem intuitiva e de fácil utilização.

A DTM, apesar da aparente preocupação com a usabilidade, deixa a desejar. Um usuário que a use pela primeira vez terá dificuldades para criar um projeto de migração de dados.

4.2. Análise comparativa das ferramentas

A Tabela 1 apresenta um resumo da análise do atendimento dos requisitos para ferramentas de migração por parte das ferramentas analisadas. Utilizamos a seguinte convenção na tabela: requisitos *Não Atendidos* são representados pelo acrônimo NA; requisitos *Parcialmente Atendidos* são representados pelo acrônimo PA; requisitos *Totalmente Atendidos* são representados pelo acrônimo A.

O MySQL Migration Toolkit e DTM são as ferramentas que atendem à menor quantidade de requisitos. As ferramentas apresentam-se de forma bem limitada, atendendo apenas às necessidades de projetos simples de migração de dados para os SGBDs mais populares.

Requisitos	SwisSQL	Jexodus	MySQL Kit	DB Convert	DTM	Data Loader
Migração envolvendo estruturas diferentes	A	A	NA	A	A	A
Migração paralela ou distribuída	NA	NA	NA	PA	NA	NA
Migração Incremental	A	A	A	A	NA	PA
Reengenharia de dados	A	A	NA	A	NA	NA
Migração envolvendo paradigmas diferentes	NA	NA	NA	NA	NA	NA
Migração entre bases de dados com modelagens diferentes	A	A	NA	NA	NA	NA
Testabilidade	PA	NA	NA	PA	PA	PA
Boa usabilidade	A	PA	A	A	PA	A

Tabela 1: Comparativo das ferramentas de migração de dados.

O JExodus atende a uma quantidade razoável de requisitos, porém fica atrás em testabilidade, por conta de não oferecer suporte a tal necessidade, além de ter uma usabilidade menor, quando comparada às demais.

O SwisSQL e DB Convert foram as ferramentas que atenderam a maior quantidade de requisitos. Todavia, as ferramentas pecam principalmente nos pontos onde as outras duas ferramentas também deixaram a desejar, tal como migração paralela e migração entre paradigmas diferentes, além de precisar evoluir mais para favorecer a testabilidade de migrações.

5. Conclusão

Este trabalho apresenta um conjunto de requisitos relacionados às ferramentas para migração de dados. O catálogo contém a materialização das necessidades, desejos e expectativas de usuários de tal tecnologia e foi construído a partir da experiência de membros da empresa Infoway em projetos de migração, aliado à pesquisa em ferramentas existentes e em trabalhos relacionados. O catálogo pode ser utilizado para facilitar o trabalho de desenvolvimento de novas ferramentas que auxiliem projetos de migração, além de servir como base para a comparação de ferramentas existentes, facilitando a identificação de produtos mais apropriados a certos projetos.

Neste trabalho avaliamos seis ferramentas de migração de dados para verificar o nível de atendimento aos requisitos propostos, além de servir como base experimental para avaliação do catálogo. Conforme comentado, existem requisitos não atendidos por

qualquer uma das ferramentas analisadas. Isso indica que ainda existe espaço para melhoria nas ferramentas existentes. Esse fato é bastante interessante, pois confirma que o catálogo pode auxiliar no desenvolvimento de novas ferramentas, além ajudar na seleção de ferramentas existentes.

Bibliografia

- Aebi, D.; Largo, R. Methods and Tools for Data Value Re-Engineering. In Applications of Databases (ADB-94), V. 819 LNCS, 1994, p. 400-411.
- Fontanette, V.; Prado, A.; Oliveira, A. L. Uma Abordagem para Migração Gradativa de Aplicações Legadas. SBES - Simpósio Brasileiro de Engenharia de Software, Brasília, DF, 2004.
- Haller, K. Data Migration Project Management and Standard Software - Experiences in Avaloq Implementation Projects. Lecture Notes in Informatics (LNI) 138, Data Warehousing Conference (DW2008), St. Gallen, Suíça, Outubro de 2008.
- Hoffmann, M.; Kuhn, N.; Weber, M.; Bittner, M. Requirements for requirements management tools. Proceedings of the IEEE International Requirements Engineering Conference (RE'04), pages 301–308, Kyoto, Japão, Setembro de 2004.
- Horstmann, J. Migration to Open Source Databases. Dissertação de Mestrado em Ciência da Computação. Technical University Berlin, Berlim. 2005.
- Howard, P.; Potter, C. Data migration in the global 2000. Relatório Técnico. United Kingdom. Bloor Research. 2007. Disponível em <http://www.techrepublic.com/>. Último acesso em janeiro de 2012.
- Lubeck, M.; Geppert, D.; Nienartowicz, K. An Overview of a Large-Scale Data Migration. Proceedings of the 20th IEEE/11th NASA Goddard Conference on Mass Storage Systems and Technologies, 2003 (MSST 2003), Abril de 2003, p. 49- 55.
- Matthes, F.; Schulz, C.; Haller, K. Testing & Quality Assurance in Data Migration Projects. 27th IEEE International Conference on Software Maintenance (ICSM), Williamsburg, Estados Unidos, Setembro de 2011.
- Ramalho J. C.; Ferreira, M.; Faria, L; Castro, R.; Boas Práticas na Migração de Repositórios: lições aprendidas com o CALM e o ARQBASE. 11º Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas: “Integração, Acesso e Valor Social“, Lisboa, Portugal, Outubro de 2012.
- Rodrigues-Neto, J.; Passos, E. JExodus: uma ferramenta de migração de dados independente de SGBD. Escola Regional de Computação Ceará, Maranhão, Piauí (ERCEMAPI), Parnaíba, PI, dezembro de 2009.
- Santos-Neto, P.; Resende, R.; Pádua, C. Requisitos para Automação de Testes de Sistemas de Informação, Anais do II Simpósio Brasileiro de Sistemas de Informação (SBSI), Florianópolis, SC, Outubro de 2005.
- Sommerville, I. Engenharia de Software. 6a. Edição. Prentice-Hall, 2003.
- Stair, R. M.; Reynolds, G. W. Princípios de Sistemas de Informação. Editora Cengage Learning, 2005.