

Lições aprendidas com a implantação de um sistema acadêmico de graduação em parceria com empresa terceirizada: Um relato de experiência

Robson Vieira do Nascimento¹, Renato de Freitas Bulcão Neto¹

¹Instituto de Informática - Universidade Federal de Goiás (UFG)
Caixa Postal 131 - 74.001-970 - Goiânia - GO - Brazil

robsonvn@ufg.br, renato@inf.ufg.br

Abstract. *This paper reports the authors' experience on the deployment process of an integrated information system for supporting undergraduate activities, named SIGAA, specially the migration of data in a Brazilian federal university. Authors describe difficulties, solutions and lessons learned both with the migration of data and the experience of working with a virtual team. That team is composed of university staff as well as employees of a third party to remotely perform the installation, configuration, deployment, customization, support, maintenance, support and consulting of the integrated information system.*

Resumo. *Este trabalho visa relatar a experiência obtida ao participar do processo de implantação, especificamente na migração de dados, do Sistema Integrado de Gestão de Atividades Acadêmicas (SIGAA), em uma universidade federal brasileira. São relatadas as dificuldades, soluções e lições aprendidas, referentes tanto ao processo de migração de dados quanto ao trabalhar com equipes geograficamente distantes, formadas por funcionários da universidade e de uma empresa terceirizada, contratada para realizar, de forma remota, a implantação do SIGAA.*

1. Introdução

O processo de migração de dados é cercado de armadilhas e desafios. Apesar de ter sua necessidade reconhecida pela governança, geralmente sua importância é minimizada, pois é vista como uma simples transferência de dados de um banco para outro, implicando em recursos e prazos incoerentes com a demanda [Oracle 2011].

A migração fica mais complexa ao terceirizar parte do processo, pois os desafios de comunicação nesse cenário são bem maiores [Oracle 2011]. Desse modo, não é uma surpresa que apenas 16% dos projetos de migração de dados são concluídos no prazo e orçamento estimado [Howard and Potter 2007].

Situação análoga ocorreu com uma universidade federal brasileira, que optou por adquirir os Sistemas Integrados de Gestão (SIG) da Universidade Federal do Rio Grande do Norte (UFRN), bem como contratar uma empresa para substituir gradativamente seus sistemas legados e migrar os respectivos dados, quando necessário. Alguns problemas foram encontrados ao decorrer desse processo de implantação e este artigo busca relatar as lições aprendidas com esses problemas.

Essas lições, podem auxiliar outras instituições federais que, futuramente, venham a optar pela aquisição do SIG. Uma vez que, a universidade federal tratada neste artigo, não foi a única Instituição Federal de Ensino Superior (IFES) a adquirir o SIG; dezenas de IFES brasileiras também optaram pelo SIG como solução de software para gestão universitária, administrativa e/ou financeira [SINFO 2015].

O presente artigo está assim organizado: a Seção 2 apresenta de forma breve o SIG; a Seção 3 descreve o cenário da implantação do SIG na referida universidade; a Seção 4 discute as justificativas para escolha do SIG; a Seção 5 detalha o processo de implantação do SIG, principalmente a migração de dados do sistema acadêmico; a Seção 6 destaca as lições aprendidas; e a Seção 7 finaliza com as conclusões obtidas com a implantação do sistema acadêmico.

2. Sistemas Integrados de Gestão - SIG

Concebido pela UFRN, o SIG é formado por um conjunto de sistemas que visam administrar informações e processos de diferentes áreas funcionais de forma integrada [SINFO 2015].

Duas grandes áreas compõem o conjunto sistêmico: (1) área meio (administrativa), que engloba o Sistema Integrado de Gestão de Patrimônio, Administração e Contratos (SIPAC) e o Sistema Integrado de Gestão e Recursos Humanos (SIGRH) e (2) área-fim (acadêmica), que o Sistemas de Integrado de Gestão de Atividades Acadêmicas (SIGAA) se enquadra. Na Figura 1 tem-se um esquema dessas áreas e seus relacionamentos.

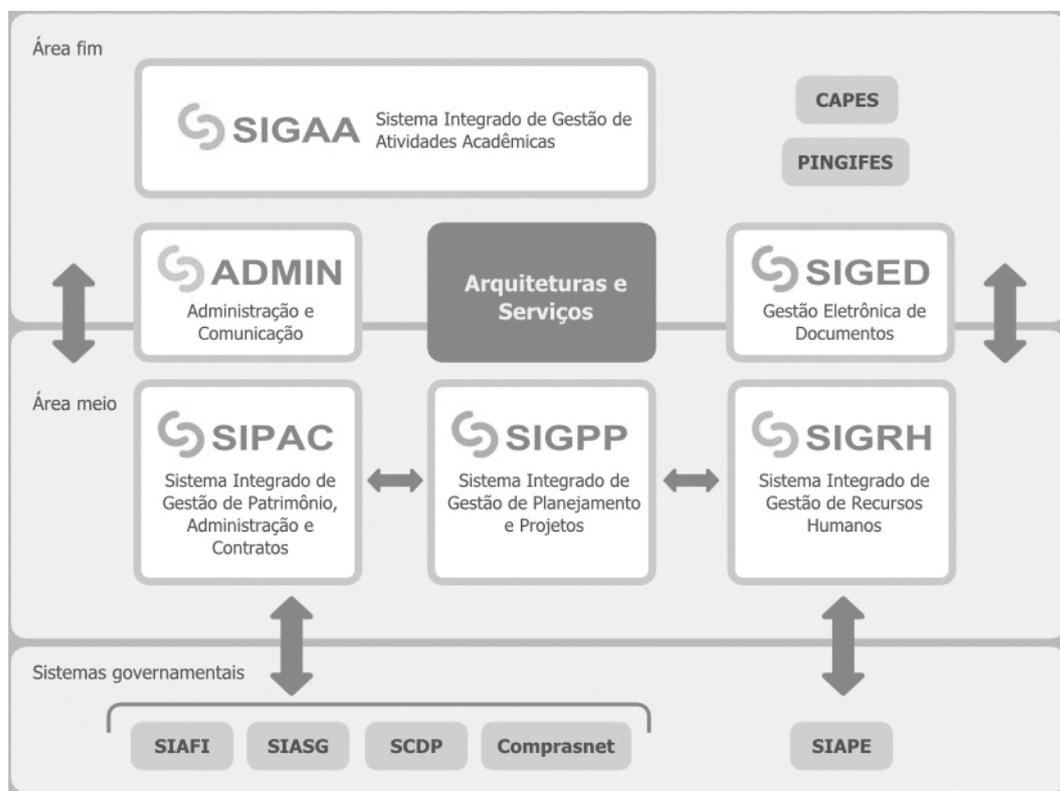


Figura 1. Diagrama de inter-relacionamento dos sistemas SIG.

3. Cenário de Implantação

Desde a sua fundação, a universidade federal de qual trata este artigo se mostra preocupada com automatização de suas atividades e acompanha a evolução tecnológica em seus sistemas. No início da década de 1970, trabalhando com *mainframe*, utilizava sistemas desenvolvidos em Fortran; posteriormente, em 1986, os sistemas foram reescritos em COBOL, que por sua vez, no novo milênio, deram espaços aos novos sistemas escritos em Oracle Forms, versão 6i, utilizando o banco de dados relacional OracleTM [CERCOMP 2014]. Isso representou uma mudança substancial no avanço tecnológico dos sistemas da universidade, colocando fim ao uso dos *mainframes* e iniciando uma era cliente-servidor.

Até então, a universidade utilizava um regime acadêmico anual, e utilizava o Sistema de Administração Acadêmica (SAA) para gerir suas atividades acadêmicas. Com advento do novo Regulamento Geral dos Cursos de Graduação (RGCG) em 2002, que previa o um regime semestral de aulas, a universidade implementou um novo sistema, o Sistema Acadêmico de Graduação (SAG), ainda em Oracle Forms, que contemplou as atualizações do novo regulamento.

A migração dos alunos do regime anual para o semestral foi gradativa. Coube as unidades acadêmicas e aos discentes a opção por migrar para o regime semestral. Essa gradatividade implicou na manutenção de dois sistemas acadêmicos, simultaneamente, em produção, para atender as demandas dos alunos que optaram por continuar no regime anual.

No SAG, o processo de matrícula do discente se dava por requisições feitas pessoalmente nas secretárias, causando um grande infortúnio aos discentes. Apesar da então recente implementação do SAG, apenas com a tecnologia utilizada no projeto não era possível a criação de um sistema de matrícula online eficiente.

Então, em 2005, foi lançado a primeira versão do sistema de matrícula online, o Matrícula Web, implementado em PHP e integrado ao SAG via banco de dados. Assim, compunha-se o universo de sistemas de graduação da universidade, ao lado do Sistema de Cadastro de Atividades Docentes (SICAD) e o SAG Web para lançamentos de notas e frequências, ambos escritos em PHP.

Ao longo de tantas implementações e refatorações, diversas migrações de dados foram realizadas. Essas migrações, somadas à diversidade de sistemas ausentes de padrões de arquitetura, fragilizaram a integridade e consistência dos dados.

Operante por mais de uma década, o SAG possuía aproximadamente quatro milhões de registros, distribuídos em cerca de trinta entidades e atendia mais de vinte mil usuários, entre discentes, docentes e técnicos.

4. Escolha do SIGAA

Em 2012, foi aprovado pelo Conselho de Ensino, Pesquisa, Extensão e Cultura (CEPEC) um novo RGCG, trazendo diversas mudanças em relação ao regulamento vigente à época, dentre elas estão: alteração da média de aprovação; mudanças nas fórmulas de cálculo do índice de prioridade; exclusão de discentes; e regulamentação da modalidade de curso a distância.

Para atender esse novo regulamento, uma série de modificações deveriam ser feitas no SAG, sendo o atendimento à regulamentação dos cursos a distância a mais impactante, devido à necessidade de distinguir as regras aplicadas a alunos presenciais dos alunos a distância, bem como fornecer um ambiente virtual para as aulas serem ministradas. Porém, a manutenibilidade do SAG deixava a desejar e a universidade não possuía contingente suficiente para implementação da demanda.

Logo, a universidade buscou alternativas para contornar a situação, como licitar a construção de um novo software ou adquirir um existente, cuja o SIGAA era uma das opções. Após uma análise minuciosa das possibilidades, a opção de adquirir o SIGAA ganhou forças, pois ele se destacava no cenário nacional por ser um sistema robusto e integrado. Então, foi criada uma comissão com representantes de várias áreas da universidade, com o propósito de analisar, as funcionalidades e deficiência do software e avaliar as customizações necessárias para o mesmo atender as demandas do novo RGCG.

Alguns integrantes da comissão destacaram que seria necessário um alto número de customizações no SIGAA para que o mesmo atenda o novo RGCG. Porém a comissão entrou em consenso por implantar praticamente todo SIG, por apresentar soluções integradas para gestão acadêmica, pessoal e administrativa, que supriria diversas carências da universidade.

Restava apenas definir o licenciamento do software e a customização do mesmo. A licença foi adquirida por esta universidade, após a celebração de um contrato de cooperação entre as partes envolvidas, onde a UFRN se exime de adaptar o SIG a realidade desta universidade, mas define a UFRN como responsável por manter a linha de base do software e transferir conhecimentos do SIG à universidade. Já em relação às customizações, esta universidade optou por contratar o serviço de uma empresa terceirizada, à época, a única homologada pela UFRN.

Os custos envolvidos com o contrato de cooperação com a UFRN somados ao custo estimado para customizar todo o SIG, mostrou-se mais vantajoso que licitar um novo sistema acadêmico para a universidade, baseado no custo médio do ponto de função do mercado e no número de pontos de função do SIGAA.

5. Implantação do SIGAA

A universidade assinou o contrato com a empresa terceirizada em abril de 2013 [DOU 2013], definindo a empresa responsável por configurar implantar, customizar, sustentar e manter os sistemas SIG por um período de quatro anos.

O processo de implantação do SIG é gradual, dividido por módulos priorizados pela universidade contratante. Cada módulo passa por validação da versão de referência e quando necessário customizações e migrações são realizadas. Para cada uma dessas etapas, necessariamente, é aberta uma Ordem de Serviço (OS) contendo o escopo, custos e prazos de entrega.

Os primeiros módulos implantados foram do Sistema Integrado de Gestão de Recursos Humanos (SIGRH) e do Sistema Integrado de Gestão da Administração e Comunicação (SIGAdmin), que são pré-requisitos para implantação do SIGAA.

Posteriormente, em setembro de 2013, foi aberta a OS de migração do módulo de Graduação, com previsão de término em dois meses. Este prazo possibilitou estipular

a primeira data de entrega do módulo em produção: fevereiro de 2014. Todavia esse prazo foi prorrogado por três vezes, o que resultou em quatro grandes migrações, cerca de setenta por cento do processo de negócio customizado e um atraso total de mais de nove meses.

O método de migração usado inicialmente assemelha-se ao ETL “*Extract, Transform and Load*”, ou em tradução livre “Extração, Transformação e Carga”. Esse método é amplamente utilizado no processo de criação de *Data Warehouse*. Ele consiste em três etapas: (1) identificação e leitura dos dados; (2) ajustes e correções; e (3) carga dos dados na base destino [Vassiliadis et al. 2002].

As atividades de migração realizadas pela empresa iniciaram com o estudo do Modelo de Entidade e Relacionamento do SAG. Posteriormente foram desenvolvidas as consultas de extração e os algoritmos de transformação e carga. A universidade se dispunha a ajudar a empresa compreender as entidades e seus relacionamentos, porém inicialmente esse recurso pouco foi explorado pela empresa. Restava à universidade esperar a conclusão da migração para validar a aplicação com os dados migrados.

Passados dois meses, a empresa disponibilizou a aplicação para testes pela primeira vez, porém a mesma constantemente apresentava erros, impossibilitando a universidade identificar se a causa do defeito era de customização, parametrização ou migração. Para solucionar os defeitos, a empresa solicitou acréscimo de um mês à data de entrega, ficando previsto para Março de 2013. Chegada a data definida, muitos erros persistiram e novamente a data de entrega foi postergada, dessa vez por três meses.

A universidade decidiu analisar as origens dos erros e descobriu que além dos erros de customização, que não serão abordadas nesse artigo, haviam erros de migração como: falhas no entendimento do conceito de algumas entidades e de seus atributos; o mapeamento errado de alguns atributos; inconsistências/incompatibilidade dos dados de origem; e consultas de extração incorretas.

Tanto os problemas de entendimento quanto de mapeamento foram solucionados por meio de vídeo conferências e outras formas de comunicação pessoal. Esses encontros serviram para solucionar dúvidas e elucidar pontos de obscuridade.

Algumas tabelas do SAG não estavam devidamente estruturadas. Existiam tabelas com colunas que faziam referências a outras tabelas, sem a definição de chave estrangeira. Também existiam tabelas com colunas que, em teoria, não deveriam permitir valores nulos. Isso possibilitou a existência de dados inconsistentes, que atrapalharam o curso da migração. Para solucionar esses defeitos, cada caso foi analisado e corrigido separadamente.

Também existiam colunas com tipo de dados diferentes entre o SAG e o SIGAA ou possuíam tamanho incompatíveis com o esperado. Muitos desses casos só foram descobertos no decorrer da migração, o que ocasionou erros, gerando retrabalho.

Para solucionar os problemas de consultas de extração incorretas, a universidade optou por rescrever as consultas de extração, devido a sua facilidade e conhecimento da base. As consultas de extração começaram a ser rescritas e na medida que elas eram criadas, eram disponibilizadas para empresa para corrigir seus algoritmos e remigrar as entidades.

Devido a migração apresentar diversos problemas, a universidade buscou conhecer o método de validação da migração usado pela empresa. Foi verificado que a empresa validava a migração quantitativamente, isto é, se houvesse “n” registros na base de origem haveria de ter “n” registros na base de destino.

Apenas com validação quantitativa não foi possível garantir que os atributos tenham sido migrados corretamente. Também existe a possibilidade do número de tuplas da entidade estar correto porém com atributos ausentes e dados corrompidos e/ou truncados.

Um dos maiores desafios em um projeto de migração de dados é certificar que o conteúdo migrado esteja íntegro, isto é, que os dados migrados não tenham sido alterados [Jindal 2011]. Apenas com o uso do ETL, não é possível garantir o sucesso almejado no processo de migração de dados. É essencial a definição de estratégias de garantia de qualidade em projetos de migração para mitigar os riscos envolvidos [Matthes et al. 2011].

A universidade verificou no contrato celebrado com a empresa que não havia definição sobre a qualidade da migração e nem os métodos de validação que seriam aplicados. Foi necessário que a universidade implementasse suas próprias rotinas de teste de migração para validar qualitativamente os dados migrados.

Inicialmente, a universidade implementou uma integração em tempo real dos bancos de origem (Oracle) e destino (Postgres), transformando dois bancos de dados físicos em um lógico, usando a ferramenta Oracle Gateway. Esta integração permite realizar em uma única consulta, a junção de duas ou mais tabelas de banco de dados distintos. A universidade fez uso desta integração para implementar consultas de validação de migração para cada entidade. Essa consulta é descrita pela seguinte fórmula:

$$(A - B) \cup (B - A)$$

, onde “A” representa a consulta de extração e “B” representa a consulta na base de destino, que simula a consulta de extração. O resultado dessa consulta é a diferença entre os dados a serem migrados e os dados migrados; o resultado esperado é nulo.

Com esse método de validação, foi possível identificar cada registro defeituoso, sendo todos os incidentes reportados de maneira formal à empresa. Em alguns casos, foi constatado que os dados migrados não refletiam as consultas fornecidas pela universidade. Ocorria que à medida que a empresa implementava o algoritmo de migração e se deparava com a necessidade de transformar e/ou tratar algum valor, a mesma efetuava as modificações nas consultas de extração, não consultando a universidade a correção das mesmas. Consequentemente, criava-se um contraste entre o que era produzido pela universidade e o resultado obtido da migração.

Para corrigir essa divergência, um funcionário da universidade foi designado para validar e corrigir todas as consultas. Também foi definido com a empresa, que qualquer necessidade de tratamento de dados deveria ser informado a universidade, para que ela tomasse as devidas decisões e correções. E caso alguma modificação necessitasse ser realizada no software de migração da empresa, a universidade estaria ciente e consideraria no momento da validação dos dados.

Até então, não existia um processo de migração definido. Era de suma importância a definição desse processo, pois diversos módulos necessitariam de migração. Foi então que a universidade buscou definir um processo de migração, identificando os atores e

suas atribuições. Esse processo define as atividades internas (estudo, auditoria, criação e validação) e superficialmente as atividades externas, executadas pela empresa, não definindo o seu *modus operandi*.

O processo foi apresentado à empresa para deixá-la a par da nova maneira de condução das atividades de migração na universidade. A partir daí, foi deixado claro as condições necessárias para que uma entidade migrada seja considerada válida. Desde então, esse é o processo utilizado para a implantação dos demais módulos.

6. Lições Aprendidas

Alterações das consultas pela empresa sem anuência da universidade

Para evitar que essa situação ocorra, foi definido que, ao invés de disponibilizar o código da consulta para a empresa, a universidade deve disponibilizar uma *View* para cada entidade em sua base de dados. Desse modo, basta a empresa realizar uma consulta simples, sem filtros, na *View*, para extrair os dados necessários.

Definições insuficientes das entidades

Um documento de definição das entidades foi criado para descrever as propriedades de uma entidade, antes que ela seja migrada. Nesse documento, a empresa define os atributos da entidade, discriminando o nome, descrição, formatação e o mapeamento dos mesmos na base de destino. Após esse documento ser preenchido, a universidade desenvolve a consulta de extração baseada nele, levando em conta os tipos de dados, tamanhos e formatos.

Ausência de métricas de validação em contrato

As métricas de validação devem ser definidas em contrato, estipulando-se as técnicas e os níveis de aceitação de serviço.

Ineficiência de testes e validações da migração

Devido à ineficiência dos métodos de validação usados inicialmente, foi percebido que além das análises quantitativas é necessária a realização de análises qualitativas. Alguns exemplos desses tipos de análises são verificação por amostragem, confronto de dados das duas bases de dados, dentre outras.

Diversas inconsistências dos dados

Uma fase de análise e validação das inconsistências e incompatibilidades foi adicionada ao processo de migração para evitar situações inesperadas referente a dados inconsistentes.

Ausência de um processo de migração de dados

Sem um processo de migração de dados definido, o projeto de migração não seguia um curso claro, o que ampliou as possibilidades de falhas. Dessa forma, foi necessário definir um processo de forma a reduzir essas possibilidades. Por questão de escopo, o processo será detalhado apenas em suas grandes fases. Essas fases estão dispostas no fluxo apresentado na Figura 2.

Estudo e documentação: Nesta fase é feito um estudo da entidade. É preenchido o documento com as projeções e os atributos necessários por parte da empresa. Posteriormente, a universidade analisa a viabilidade da requisição e verifica se existe algum dado relevante que não foi solicitado.

Análise e correção de inconsistências: Após a definição dos atributos que serão migrados, é feita uma análise da consistência dos mesmos. Caso alguma inconsistência seja encontrada, ela é devidamente corrigida.

Elaboração de consultas: Momento onde as consultas de extração e validação são criadas e validadas.

Migração: Os dados são carregados na base de destino.

Validação: As rotinas de teste são executadas e os defeitos encontrados são documentados e corrigidos, até que a migração receba o aceite pela universidade.



Figura 2. Fluxo do processo de migração elaborado pela universidade.

7. Conclusões

Ao longo deste artigo, fez-se um relato da experiência dos autores com o processo de implantação de um sistema de informação para atividades acadêmicas de Graduação em uma universidade federal brasileira. Destaque maior foi dado à atividade de migração de dados, haja vista que a universidade já contava com uma grande massa de dados de seus sistemas em produção, bem como dos legados.

Vários problemas foram enfrentados na implantação do sistema acadêmico, porém após a equipe de implantação da universidade definir o processo de migração e o método de validação de dados migrados paulatinamente os efeitos positivos começaram a surgir, colaborando com a melhoria da qualidade das entregas e consistência da aplicação, até a migração ser aprovada em novembro de 2014. Desde então, até a presente data, o SIGAA mantém-se em produção e está em contínua evolução, apesar das tempestividades, atendendo de forma satisfatória a comunidade acadêmica da universidade.

Referências

- CERCOMP (2014). Histórico do CERCOMP. <http://cercomp.ufg.br/p/191-historico-do-cercomp>, Último acesso em Agosto de 2015.
- DOU (2013). Diário Oficial da União (DOU). In *Seção 3*, page 57. Imprensa Nacional.
- Howard, P. and Potter, C. (2007). Data migration in the global 2000 - research, forecasts and survey results. Technical report, Bloor Research.
- Jindal, S. (2011). Effective testing quality assurance in data migration projects. Technical report, LT Infotech.
- Matthes, F., Schulzl, C., and Haller, K. (2011). Testing & quality assurance in data migration projects. *IEEE International Conference on Software Maintenance*.
- Oracle (2011). Successful data migration. Technical report, Oracle Corporation.
- SINFO (2015). Sobre a cooperação técnica. <http://www.portalcooperacao.info.ufrn.br/pagina.php?a=sobre>, Último acesso em Agosto de 2015.
- Vassiliadis, P., Simitsis, A., and Skiadopoulou, S. (2002). Conceptual modeling for ETL processes.