

HIT – Aspekte einer Migration von Einzelplatzlösungen in webbasierte Client-Server-Systeme

Sum, R.
MAN Nutzfahrzeuge AG, Abteilung TDB
Dachauer Straße 667, 80995 München
richard.sum@man.eu

Loos, A.
pro et con GmbH
Annaberger Straße 240, 09125 Chemnitz
andreas.loos@proetcon.de

Abstract: HIT (Harmonisierung und Integration des MAN-TDB-Datenbestandes) ist ein Migrationsprojekt, welches abteilungsspezifische Einzelplatzlösungen in ein webbasiertes System konvertiert. Im Vergleich zu klassischen Migrationsprojekten beinhaltet HIT einen signifikanten Anteil neuentwickelter Komponenten sowie Schnittstellen zu externen Prozessen. Im Rahmen des Projektes wurden spezielle Migrationstools für konkrete Anwendungsfälle entwickelt, auf den Einsatz von Standardsoftware wurde verzichtet.

1 Motivation und historische Entwicklung

Die Abteilung „Technische Dienstleistung und Beratung“ der MAN Nutzfahrzeuge AG (MAN-TDB) erstellt Bestätigungen und Genehmigungen für Umbau resp. Umrüstung von Nutzfahrzeugen. Die entstehenden Dokumente werden weltweit an Werkstätten, Service-niederlassungen und Aufbauhersteller versendet. Zu Beginn der 90er Jahre erhöhte sich die Zahl der entstehenden Vorgänge drastisch, so daß eine computergestützte Verwaltung unumgänglich wurde. Es entstanden weitgehend Einzelplatzlösungen, die essentiell den innerhalb MAN-TDB existierenden Arbeitsgruppen zugeordnet waren. I.d.R. wurden dabei Microsoft-Office-Produkte eingesetzt.

Durch die Dynamisierung der Märkte, die minimale Reaktionszeiten erfordern, durch das Ansteigen der Vorgangszahlen und die steigende Variantenvielfalt der MAN-Nutzfahrzeuge erwies sich die Lösung Ende der 90er Jahre als nicht mehr tragfähig. Es wurde ein typisches Dilemma aller Einzelplatzlösungen offensichtlich: Datenmodelle und Datenstände, z.B. für eine Kundenverwaltung, die allen MAN-TDB-Anwendungen inhärent sind und damit eine redundanzfreie Darstellung erfordern, drifteten im Laufe der Zeit zwischen den einzelnen Anwendungen auseinander. Zudem existierte keine Möglichkeit, die bei MAN in großer Zahl existierenden zentralen Prozesse zu integrieren, so daß ein Bearbeiter eine Vielzahl heterogener Programme auf seinem PC benutzen mußte, um über „Copy-and-Paste“ Daten auszutauschen.

Zur Problemlösung wurden unterschiedliche Strategien diskutiert - einerseits in Richtung etablierter Standardsoftware und andererseits in Richtung abteilungsspezifischer Webappli-kation auf Basis einer Abteilungssoftware, mit welcher der gesamte MAN-TDB-interne

Workflow modelliert werden kann. Grund für die Entscheidung zugunsten der zweiten Variante war, daß die Anwendung damit hochgradig flexibel und skalierbar wird. Eine schnelle Reaktion auf modifizierte Situationen ist gegeben. Ebenso ist ein feingranulares Design möglich, mit welchem auch detaillierte Anforderungen der Abteilung umgesetzt werden können. Diesen Ansprüchen werden Standardprodukte nicht in jedem Fall gerecht.

In [Su05] wurde bereits überblicksmäßig über HIT berichtet. Der vorliegende Artikel beschreibt das Projekt aus technischer Sicht im Fokus der Datenmigration.

2 Randbedingungen der Migration

Für HIT wurden Kernziele definiert:

1. Die Geschäftsprozesse werden vollständig neu modelliert. Auf Basis eines relationalen Entwurfs wird ein Datenmodell erstellt, welches einerseits die bestehende Struktur widerspiegelt und andererseits zukunftssicher genug ist, um wechselnden Anforderungen des Marktes gewachsen zu sein.
2. Der existierende Datenbestand muß automatisiert migriert werden. Kern ist dabei eine Konsolidierung, um vielfältige Redundanzen zu beseitigen. Damit einher geht eine Normierung von Notationsformen. Unterschiedliche Schreibweisen ein- und desselben Sachverhaltes z.B. müssen während der Migration analysiert und mit hohem Automatisierungsgrad migriert werden. Als Komponente der Qualitätssicherung ist im Zielsystem die Einhaltung dieser Normen zu prüfen. Basis sind unternehmensweite Standards ebenso wie DIN-Normen. HIT beinhaltet Kernfunktionalitäten wie u.a. die Kundenverwaltung, welche in allen weiteren HIT-Modulen Verwendung finden. Die Modellierung erfolgt arbeitsgruppenspezifisch. Die Module werden sequentiell entwickelt und zu unterschiedlichen Zeitpunkten in die Produktion überführt, um die Arbeitsfähigkeit der Abteilung während der Einführungsphase zu sichern. Die Kopplung der Module erfolgt durch einen definierten Workflow, welcher typische Arbeitsprozesse und Schnittstellen zwischen den Arbeitsgruppen abbildet.
3. Die Datenmigration schafft die Voraussetzung für eine automatisierte Archivierung von Dokumenten. Damit werden Anforderungen an die Revisionssicherheit und die gesetzlich vorgeschriebenen Aufbewahrungsfristen für Dokumente erfüllt.
4. Die Migration des Datenbestandes beinhaltet eine Schemamigration. Das im Ausgangssystem definierte Schema entsprach nicht dem relationalen Schema, welches im Zielsystem Verwendung finden sollte.
5. HIT muß zu einer spürbaren Entlastung der Mitarbeiter führen. Routinearbeiten, z.B. Verwaltungstätigkeiten, werden auf ein Minimum reduziert. HIT muß mit minimalem Schulungsaufwand erlernbar und nahezu intuitiv bedienbar sein. Das System wird als hochverfügbar eingestuft, um eine permanente Nutzung zu garantieren. HIT wird in späteren Ausbaustufen MAN-Nutzern weltweit an verschiedenen

Produktionsstandorten zur Verfügung stehen. Daher wird es als Webapplikation ausgeprägt.

6. HIT integriert systemübergreifende Prozesse wie Fakturierung, revisionssichere Langzeitarchivierung, Stammdatenimporte aus Mainframe-Anwendungen uvm.
7. HIT muß so flexibel sein, daß Funktionalitäten in kürzester Zeit modifiziert oder nachgerüstet werden können. Eine vollständig neue Funktionalität z.B. stellt die automatische Dokumentengenerierung in unterschiedlichen Fremdsprachen dar. In der ersten Ausbaustufe werden dabei 7 Sprachen unterstützt.

3 Spezielle Migrationsaspekte

Das Projekt bei MAN-TDB unterscheidet sich von klassischen Migrationsprojekten:

An Migrationsprojekten sind in der Regel IT-Abteilungen der Unternehmen und externe Unternehmen, welche die Migration durchführen, beteiligt. Das Personal rekrutiert sich fast ausschließlich aus IT-Fachleuten. MAN-TDB besitzt spezialisierte Mitarbeiter auf dem Gebiet der Kraftfahrzeugtechnik, der Elektrotechnik und des Maschinenbaus. Aus diesem Grund mußte zunächst ein einheitlicher Begriffsapparat geschaffen werden, auf dessen Basis die Kommunikation zwischen MAN-TDB und dem mit der Migration beauftragten Unternehmen möglich ist.

Migrationsprozesse überführen einen Ausgangszustand TA in einen Zielzustand TZ , wobei die anwenderseitige Funktionalität von TZ identisch mit der in TA ist. Diese Vorgehensweise wird gewählt, um die ohnehin schon risikobehaftete Migration nicht mit zusätzlichen Risiken neuentwickelter Komponenten zu erschweren. HIT beinhaltet sowohl Migration als auch Neuentwicklung.

Da wesentliche Abschnitte der Migration aus Zeitgründen vollautomatisch erfolgen mußten, waren Migrationstools unumgänglich. Damit stand wiederum die Frage nach der Nutzung von Standard- oder projektbezogener Spezialsoftware. Auch hier fiel die Entscheidung auf die zweite Variante, da Migrationssoftware einerseits nur für einen eng begrenzten Zeitraum benötigt wird und mit minimalem Userinterface ausgestattet sein muß, andererseits aber detailliert parametrierbar sein soll, um konkreten Projektanforderungen zu entsprechen.

Eine wesentliche HIT-Komponente ist die Integration externer, heterogener Systeme, die über Datenschnittstellen miteinander gekoppelt sind. Die Modellierung des Datentransfers ist damit ein wichtiges Designkriterium. Die Analyse externer Prozesse zeigte, daß Daten teilweise nur in proprietären Formaten geliefert werden können. Eine allgemeingültige Schnittstelle existiert nicht. Aus diesem Grund wurde eine Datenabstraktionsschicht in die Applikation integriert, die unterschiedliche Datenformate in XML-Formate transformiert. Den Funktionalitäten der Geschäftslogik bietet sich damit ein syntaktisch homogener Datenstrom, der durch gängige XML-Parser analysiert wird. Die XML-Funktionalität wird neben der Datenabstraktionsschicht weiterhin im Userinterface genutzt. HIT arbeitet mit Ajax- (Asynchron Javascript and XML) Methoden und bietet damit eine performante

Schnittstelle, in der XML-Datenströme zwischen Server und Client ausgetauscht werden. Abbildung 1 zeigt eine schematische Darstellung der HIT-Applikation:

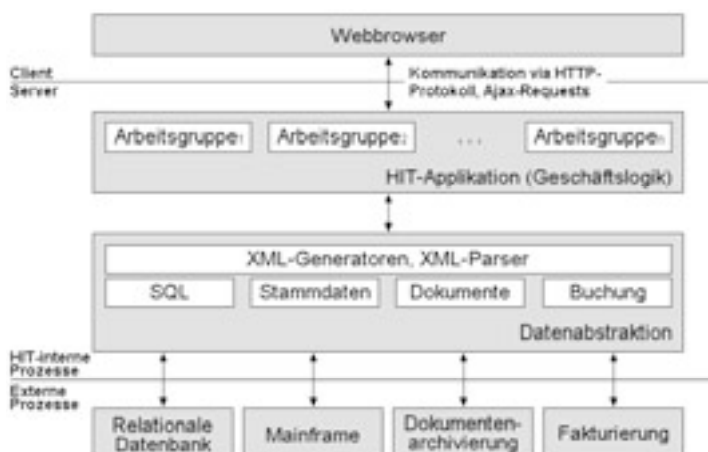


Abbildung 1: Schematische Darstellung der HIT-Applikation

Die Datenmigration mußte vollautomatisch erfolgen. Mit der Umstellung auf HIT wurden alle relevanten Daten des Altsystems in eine relationale Datenbank übernommen. Der Konsolidierung des Datenbestandes wurde in diesem Zusammenhang besondere Aufmerksamkeit gewidmet. Dabei kamen u.a. folgende Strategien zum Einsatz:

Schemamapping: Bedingt durch Differenzen zwischen Ausgangs- und Zielschema wurde ein Schemamapping durchgeführt, welches u.a. das Mapping korrespondierender Attribute [LN07] und die Überführung des Schemas in die 3. Normalform beinhaltete. Die formale Beschreibung des Schemamappings (Mappingregeln) erfolgte manuell. Tools kamen hier nicht zum Einsatz. In dieser Phase wurden MAN-TDB-Mitarbeiter in die Entwicklung integriert, was zu einer intensiven Nutzung vorhandenen Know-hows und zu erhöhter Lösungsakzeptanz führte. Der Migrationsprozeß erfolgte auf Basis deskriptiver Mappingregeln unter Einsatz eines speziell für diese Aufgabe entwickelten Tools.

Normenparser: Unter diesem Begriff wurden Tools auf der Basis regulärer und kontextfreier Grammatiken [LZ06] entwickelt, die alle in HIT-Datenbanken enthaltenen Attribute, deren Format durch deutsche bzw. internationale Normen oder unternehmensinterne Standards syntaktisch beschrieben werden kann, analysieren. Typische Beispiele sind Telefon- und Faxnummern oder Postleitzahlen. Die Software unterscheidet sich von anderen Migrationstools dadurch, daß die enthaltenen Algorithmen auch im HIT-System selbst zur Anwendung kommen, um die syntaktische Korrektheit neu erfaßter Daten zu verifizieren.

Editierabstand: Der Editierabstand [Le65] beschreibt die Zahl notwendiger Editieroperationen, um eine Zeichenkette S_1 in eine Zeichenkette S_2 zu überführen. Auf diese Weise wird die Ähnlichkeit zweier Zeichenketten ermittelt. Ist der Editierabstand von S_1 und S_2 gering, kann in gewissen Situationen auf die Identität von S_1 und S_2 geschlossen werden. Differenzen treten in diesem Zusammenhang u.a. durch Schreibfehler auf. Durch dieses Vorgehen wurden ca. 80 % der Inkonsistenzen im Datenbestand ermittelt und beseitigt.

Manuelle Nacharbeiten waren notwendig. Diese wurden nach der Migration im Zielsystem ausgeführt.

Auf Basis der gewonnenen Ergebnisse wurde ein Regelsystem erstellt, welches die Transformation fehlerhafter in korrekte Daten beschreibt.

4 Erfahrungen und Ausblick

HIT arbeitet seit 2005 stabil. Es wurde das Basissystem einschließlich aller arbeitsgruppenspezifischer Applikationen entwickelt. Das HIT-Webinterface bietet dem Bearbeiter einen homogenen Zugang. Daten werden zentral und redundanzfrei verwaltet. HIT erfährt eine hohe Akzeptanz, was auf die intuitive Bedienung, die Reduktion administrativen Overheads und die aktive Einbeziehung aller HIT-Anwender während der Migrations- und Entwicklungsphase zurückzuführen ist. Die XML-Datenabstraktionsschicht ermöglicht die Definition von Schnittstellen zu weiteren, externen Prozessen, ohne die interne Geschäftslogik von HIT modifizieren zu müssen.

Durch den massiven Einsatz von Ajax konnte die Applikation so optimiert werden, daß Performanceverluste, die in Webanwendungen durch den Transfer großer Datenmengen zwischen Client und Server entstehen, kompensiert wurden.

Bilanzierend ist festzustellen, daß sich die gewählten Strategien als robust und tragfähig erwiesen haben. Das betrifft insbesondere die Festlegung, keine Standard-Tools einzusetzen, sondern applikationsspezifische Migrationssoftware zu entwickeln. Migrationen sind größtenteils projektspezifisch. Tools werden nur kurze Zeit benötigt.

Verbesserungen sind im Bereich der Datenkonsistenz möglich. Ca. 20 % manuelle Nacharbeiten sind nur in Projekten akzeptabel, in denen mit überschaubaren Datenmengen gearbeitet wird.

Die vollständige Modellierung des MAN-TDB-Workflows ist eine neue Phase im HIT-Projekt. Hier liegt das Augenmerk auf der interdisziplinären Kopplung der einzelnen Arbeitsgruppen.

Literaturverzeichnis

- [Le65] Levenshtein, V.I.: Binary codes capable of correcting deletions, insertions and reversals. In: Doklady Akademii Nauk SSSR, 163(4) S. 845-848, 1965 (Russisch).
- [LN07] Leser, U.; Naumann, F.: Informationsintegration. dpunkt, 2007, ISBN 3-89864-400-6.
- [LZ06] Loos, A.; Zimmermann, Y.: Das HIT-Projekt. Technische Dokumentation, Version 2.3, pro et con GmbH, 2006.
- [Su05] Sum, R.: Das Migrationsprojekt "Harmonisierung und Integration von Datenbanken bei MAN/TDB". Fachberichte Informatik, Universität Koblenz-Landau, 15/2005.