

Innovativer DWH Blueprint – BI Konzeption bei DB Schenker Rail

**Dominic Marx / DB Schenker Rail / Mainz
Knut Veltjens / CGI / Frankfurt**

Schlüsselworte

Data Warehouse Handbuch / Blueprint; Oracle Exadata; Oracle Data Integrator (ODI); DataVault; Aufbau BICC; SAP Business Objects, GreenIT; CO² Einsparungen; Corporate DWH; Best Practice Entwicklung & BICC; BI Framework; Data to Diamonds; Sizing DWH Infrastruktur

Einleitung

Das Projekt hatte zum Ziel, das BICC (Business Intelligence Competence Center) der DB Schenker Rail inhaltlich zukunftsfähig für die anspruchsvollen Herausforderungen im internationalen Konzernkontext aufzustellen und die Basis für ein Data Warehouse mit hohen Performanceanforderungen zu schaffen. In Zusammenarbeit mit dem Beratungshaus CGI wurde neben dem Aufbau von BICC Strukturen ein Blueprint zur Data Warehouse Entwicklung erstellt, der zukünftig die Voraussetzungen für eine international verteilte Zusammenarbeit und eine Skalierbarkeit der Umsetzungsaktivitäten bei gleichbleibend hoher Qualität bildet. Das CGI Data to Diamonds Business Intelligence Framework lieferte für das Projekt Best-Practice-Vorgehen für Aktivitäten, Modelle und Technologien, die auf etablierten BI-Architekturen und Verfahren basieren und die fachlichen und technischen Anforderungen in den Themenfeldern BI-Strategie, BI-Spezifikation, BI-Entwicklung und BI-Betrieb ebenso abdecken wie Projekt- und Organisationserfordernisse.

Erste Entscheidungen

Im Zuge des Projekts wurden zudem richtungsweisende Entscheidungen der im BICC einzusetzenden Hard- und Softwaretechnologien getroffen. Neben dem Reporting- und Analyse-Tools und der Datenbanktechnologie wurden auch die Komponenten zur Datenbewirtschaftung und Modellierung untersucht und festgelegt. Für die Modellierung des Core DWH wurde im Blueprint der DataVault Ansatz festgeschrieben.

Heute bilden der Oracle Data Integrator (ODI) als Datenbewirtschaftungstool, die Oracle Datenbank als zentrale Warehouse-Datenhaltung sowie die Oracle Exadata Appliance die technologische Basis des Corporate Data Warehouse bei DB Schenker Rail. Für die Modellierung des Datenmodells wird der SAP Sybase PowerDesigner eingesetzt. Als Frontend wird SAP BusinessObjects zur Interaktion mit dem Anwender (Analyse und Reporting) eingesetzt.

Blueprint

Das im Projekt entstandene Blueprint Dokument (Handbuch) enthält die kompletten Informationen und Vorgaben zur Entwicklung eines Core Data Warehouse (DWH) inkl. Grundlagen, Regeln & Templates.

- Grundlagen im Handbuch beschreiben detailliert u. a. die Ausgangssituation und Zielarchitektur aber auch fachliche Einführungen in weitere Data Warehouse- und Business Intelligence betreffende Thematiken.
- Herzstück des Handbuches ist die Definition und Ableitung eines ETL Frameworks zur täglichen Arbeit mit dem ODI.

- Die definierten Regeln sind Vorgaben zur Umsetzung von Projekten und Aufgabenstellungen im jeweiligen Kontext (Import / Export, Modellierung & ETL, Reporting).
- Ergänzt werden die beschriebenen Informationen durch Templates aus der Praxis zur Dokumentation von fachlichen und technischen Spezifikationen (z. B. Datenquellen-Template, Kennzahlen-Template).

Das entstandene - mehr als 300 Seiten starke - Handbuch inkl. der Templates für die tägliche Projektarbeit - von der Anforderungsanalyse über die Entwicklung bis zur erfolgreichen Überführung in den Betrieb - behandelt unter anderem die Themen:

- Prozesse & Organisation im BICC,
- Data Warehouse Architektur,
- Datenmodellierung,
- Datenhaltung,
- Datenbewirtschaftung,
- Reporting & Analyse,
- Setup & Sizing der Hardwareinfrastruktur.

Die für ein Core DWH essentiellen Bestandteile sind bis zur Ebene der technischen Anwendung in diesem Handbuch beschrieben. Das Know-how wurde in über 20 Meet-the-Expert Veranstaltungen an die Mitarbeiter des BICC vermittelt.

Infrastruktur / GreenIT

Nach der Entscheidung über die einzusetzenden Softwaretechnologien sind in einem umfangreichen Business Case unterschiedliche Varianten für ein Hardware Sizing ausgearbeitet worden. Hier ist die Entscheidung für eine Oracle Exadata Appliance getroffen worden. Neben technischen Leistungskriterien spielten auch Aspekte der GreenIT eine ausschlaggebende Rolle: Für seine Kunden hat DB Schenker Lösungen entlang der gesamten Lieferkette entwickelt, mit denen bei Gütertransporten weniger CO² ausgestoßen wird. Im Projekt wurde dieser Gedanke konsequent weitergedacht. Ergebnis ist eine Vermeidung von knapp 70% CO² Ausstoß durch verminderten Stromverbrauch im Vergleich zu einer herkömmlichen DWH-Infrastruktur. Im Rechenzentrum entspricht dies mehr als 1.000 Tonnen CO²-Einsparung über einen Zeitraum von 10 Jahren. Von der Kostenseite wird ein Volumen von knapp 0,6 Mio. Euro über diesen Zeitraum eingespart (zu Preisen von 2013).

BICC / Wandel um 180°

Im BICC fand eine Umorganisation der Verantwortlichkeiten und Neudefinition der Rollen und Aufgaben statt. Ehemals anwendungs- oder themenbezogene Verantwortung ist heute nach den folgenden Bereichen gegliedert:

- Architektur/Datenmodell
- Spezifikation
- ETL / Coding
- Frontend-Entwicklung

Konkret bedeutet dies eine Abkehr von Verantwortung für einzelnen Fachanwendungen (IT Systeme z.B. Data Marts) hin zu einer Spezialisierung in einem der oben genannten Felder. Hierdurch wird eine deutliche Stabilisierung der Abteilung erreicht.

Technisches Know-how ist fast automatisch mehrfach vorhanden und durch klare Vorgaben an die Spezifikation von Anforderungen wird eine langfristig ausgelegte Sicherung des Prozesswissens vorangetrieben. Hierdurch wird aktiv dem Aufbau von „Herrschaftswissen“ vorgebeugt.

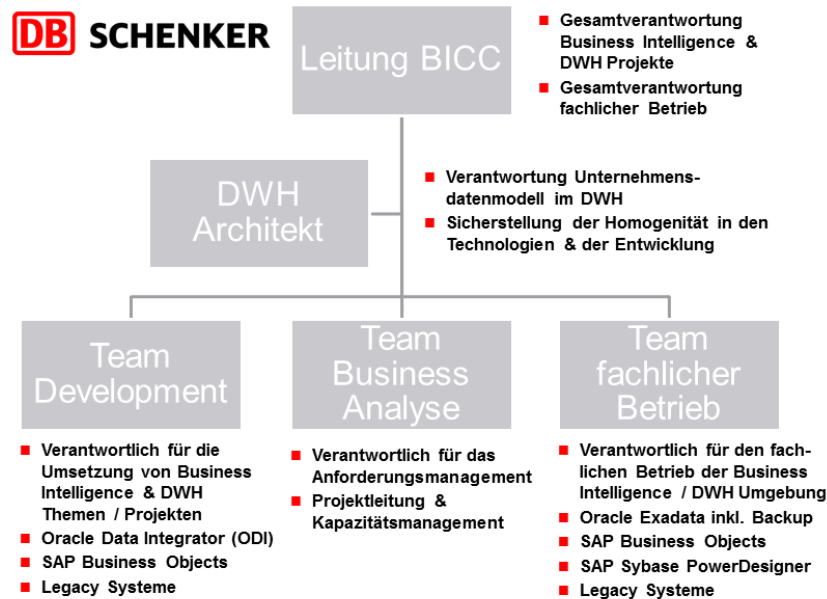


Abb. 1: Organisation DB Schenker BICC

Aufgrund der Größe der Organisation und den umfangreichen Änderungen an den Arbeitsabläufen wurde von Beginn an ein klar strukturierter, einfacher Aufbau für die Gestaltung des BICC gewählt. Im Wesentlichen kommt das BICC heute mit fünf Verantwortungsbereichen aus.

DataVault

Als Modellierungsmethodik für das DWH wurde Data Vault gewählt. Die Methode stellt einen Hybrid-Ansatz aus 3NF und Star Schema dar und bietet neben einer impliziten Historisierung und Versionierung eine hohe Flexibilität bei Änderungen im Datenmodell. Bei der Modellierung mit DataVault stehen die Geschäftsdaten im Fokus, indem zunächst Geschäftsobjekte (bspw. Kunde, Auftrag, Zug oder Verladestelle) identifiziert werden. Vorteile liegen zudem in der Erreichung einer hohen Ladeperformance sowie in der Gestaltung eines flexiblen Systems, das sich schnell und einfach anpassen lässt, sowie skalierbar, nachvollziehbar und auditierbar ist.

1. Projekt erfolgreich

Im ersten, basierend auf dem DWH Blueprint und den neuen Arbeitsabläufen, umgesetzten DWH-Projekt wurde das interne Rechnungswesen der DB Schenker in die zukünftige DWH Umgebung portiert. Kernstück der verarbeiteten Daten bilden Buchungssätze, die als Verrechnungsvorgänge in einem einheitlichen Modell abgebildet sind. Mit dem Projekt wurden gleichzeitig die Stammdaten des Rechnungswesens in eine einheitliche Form gebracht.

Auf der Reporting Ebene wurden verschiedene Standardansichten und zusätzlich Self-Service Bereiche zur Verfügung gestellt, die Daten aus unterschiedlichen Perspektiven anzeigen können. Hier sind die Daten in aufgearbeiteter Form zu den Verrechnungsvorgängen abrufbar. Unter anderem wurden Buchungssätze über Hierarchien verdichtet, sowie Ist- und Soll-Werte in selektiver und kumulierter Form gegenübergestellt.

Das interne Rechnungswesen hat vielfältige fachliche Bezüge zu anderen komplexen Fachdomänen (Beauftragung, Abrechnung, Produktionsdurchführung) und bildet somit die Basis für zukünftige Projekte des BICC. Übersicht und Mengenangaben zum Umfang des Projekts:

- Grundlagen des zukünftigen Unternehmensdatenmodells auf Basis DataVault wurden geschaffen – viele weitere Projekte werden davon profitieren.
- Über 200 Seiten technische Spezifikation zur Vorbereitung der Umsetzung wurden erstellt.
- Über 300 ETL Fragmente wurden nach Vorgaben des DWH Blueprints erstellt.
- Apex Masken zur direkten Erfassung von Stammdaten durch den Fachbereich wurden aufgebaut.
- Über 20 Stammdatentabellen können so - unter Vermeidung diverser Fehlerquellen - in einem einheitlichen Format gepflegt werden.
- **Ergebnis:** Ein zufriedener Kunde – das Projekt mit C-Level Visibilität wurde erfolgreich ausgerollt und steht nun allen Landesgesellschaften zur Verfügung.

Lessons Learned

Nachdem seit einem Jahr die Arbeitsergebnisse des Blueprints intensiv genutzt werden und mit diesem Rahmen erste Projekte erfolgreich realisiert wurden, steht fest, dass der eingeschlagene Weg, der Richtige war. Das erste Projekt hat zwar etwas länger gedauert als geplant (statt 4 Monaten wurden ca. 6 Monate benötigt; aufgrund kleinerer Änderungen in den Anforderungen und der parallelen Umorganisation im BICC). Nutzer konnten aber aufgrund des sehr strukturieren Ansatzes und der entsprechenden Dokumentation für den Fachbereich bereits während der Entwicklung die Qualität der Lösung beurteilen.

Das nun im BICC etablierte Vorgehen bei Projekten wurde von den unterschiedlichen Fachabteilungen bereits adaptiert und heute ist das BICC in der Lage parallel an mehreren Projekten gleichzeitig zu arbeiten, ohne Qualitätseinschränkungen in Kauf nehmen zu müssen.

Die Einführung der neuen Arbeitsweise im BICC wurde in einem Übergangszeitraum von ca. 4 – 6 Monaten erreicht. Hierbei wurde auf ein intensives Coaching der einzelnen Mitarbeiter durch erfahrene Berater vertraut.

Im Projekt ist ein innovativer BI-Ansatz mit State-of-the-Art Technologien und CO² optimierter Infrastruktur entstanden. Dies ist unter Berücksichtigung der Empfehlungen und enger Zusammenarbeit mit dem externen Consultingpartner CGI und dem internen IT-Dienstleister im Deutsche Bahn Konzern gelungen. Mit Vor- und Nacharbeiten wurden diese Ergebnisse über einen Zeitraum von ca. 12 Monaten erzielt.

Bereits nach ca. 6 Monaten konnte parallel mit der Vorbereitung erster Projektansätze, dem Aufbau der zentralen DWH-Plattform sowie der Reorganisation hin zu einem BICC begonnen werden. Nach weiteren 3 Monaten wurde das erste, komplett nach dem neuen Ansatz durchzuführende, Projekt gestartet.

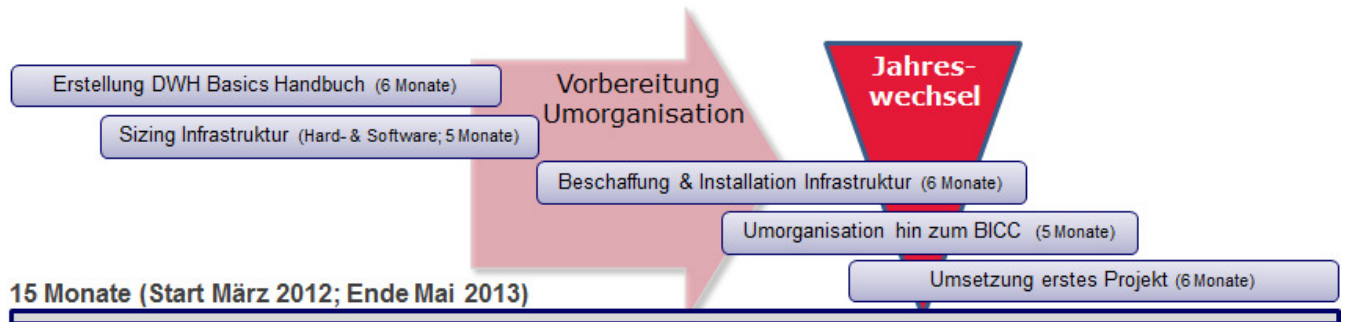


Abb. 2: Timeline Entwicklung und Einführung des innovativen DWH Blueprints

Knapp 15 Monaten nach dem Start der Aktivitäten für einen innovativen DWH Blueprint wurde das erste Projekt durch den Fachbereich abgenommen. Der eigentliche Transformationsprozess ist damit abgeschlossen. Heute finden auf dieser Basis kontinuierliche Verbesserungen der Einzelbereiche statt.

Fazit: Wenn einer internen Business Intelligence Abteilung der Spielraum und die Möglichkeit zum Aufbau einer innovativen Entwicklungsmethode in Kombination mit entsprechender Infrastruktur gegeben wird, kann auch mit überschaubaren Mitteln und mit einer begrenzten Anzahl an Mitarbeitern ein leistungsfähiges BICC aufgebaut werden.

Kontaktadresse:

Knut Veltjens

CGI Information Systems and Management Consultants (Deutschland) GmbH

Am Limespark 2

D- 65843 Sulzbach (Taunus)

Telefon: +49 (0) 178 / 8855774

E-Mail knut.veltjens@cgi.com

Internet: www.cgi.com