

Kapitel 17: Föderierte Datenbanken, Workflow-Management und Data-Warehouses

In großen Unternehmen existiert oft ein breites Spektrum, weitgehend unabhängig voneinander entwickelter und auf unterschiedlichen Softwarekomponenten basierender, datenbankgestützter Anwendungssysteme, die mehr oder weniger „Insellösungen“ darstellen. Mehr und mehr entsteht jedoch der Bedarf zur Integration solcher Insellösungen, um betriebliche Abläufe globaler optimieren zu können und Marketing- wie auch unternehmensstrategische Entscheidungen besser vorbereiten zu können. Für die sich hiermit stellende Problemstellung der Integration heterogener, strukturierter, durch Schemata beschriebener Datenquellen wird generell eine sogenannte *Wrapper-Mediator-Architektur* verfolgt, die aufbaut auf

- je einem *Wrapper* pro Datenquelle, der die eigentliche Datenquelle (z.B. eine relationale Datenbank) kapselt und deren Daten oder Operationen in aufbereiteter, zur Integration vorbereiteter Form bereitstellt, und
- mindestens einem *Mediator* (Broker, Föderations-Integrator), der die eigentliche Integration der von den Wrappern bereitgestellten Daten oder Operationen leistet und den Clients integrierte Daten und/oder Operationen anbietet.

Je nach „Tiefe“ der Integration und Aufgabenteilung zwischen Wrappern, Mediator und ggf. den Clients ergeben sich verschiedene Familien von Integrationsarchitekturen für strukturierte Daten, die im folgenden skizziert werden.

17.1 Föderierte Datenbanken

Bei der Architektur einer aus mehreren *Komponenten-Datenbanken* bestehenden *Datenbank-Föderation* stellen Wrapper – hier auch *Datenbank-Adapter* genannt – ein aufbereitetes, zur Integration vorbereitetes Schema der jeweiligen Komponenten-Datenbanken zur Verfügung, lassen aber alle Daten unter der Kontrolle des jeweiligen Komponentensystems.

Häufig wird nur ein Teil der jeweiligen zur Integration benötigt bzw. freigegeben; daher wird das vom Wrapper bereitgestellte Schema häufig auch *Export-Schema* genannt. Dieses entspricht quasi einer View auf die eigentlichen Daten des Komponenten-Systems. Der Wrapper muß daher zusätzlich eine Abbildung dieser View auf das Original-Schema der Komponenten-Datenbank, im wesentlichen also eine entsprechende Familie vordefinierter Queries, beinhalten.

Der Mediator – hier auch *Föderations-Server* genannt – integriert die exportierten „lokalen“ Schemata der Komponenten in ein globales Schema, das in einer Meta-Datenbank abgelegt wird, und unterstützt Queries und Update-Operationen auf den gemäß dieses globalen Schemas virtuell integrierten Daten. Dazu sind globale Queries in Subqueries zu zerlegen, die jeweils auf den exportierten Views ausgeführt werden, indem die jeweiligen Datenbank-Adapter die Viewdefinitionen in den Subqueries substituieren und die resultierenden Queries ausführen. Der Föderations-Server übernimmt auch die Integration der von den Komponenten-Systemen bzw. Datenbank-Adaptoren gelieferten Teilresultate zu einem globalen Resultat. Um temporäre Zwischenresultate weiterverarbeiten (z.B. aggregieren) zu können, kann dem Föderations-Server ggf. eine spezielle Föderations-Datenbank zur Verfügung stehen.