# Wenn Applikationsserver nicht ausreichen

Martin Schulze

Senior Systems Engineer, Object Design (mschulze.objectdesign.com)

Enterprise Java Beans (EJB) ist der Standard für die Entwicklung skalierbarer E-Business-Applikationen in Bereichen, die von E-Commerce über Supply Chain Management und Customer Relationship Management (CRM) bis hin zu unternehmensweiten Informationssystemen, Anwendungsportalen und anderen Anwendungsgebieten reichen. Wenngleich Applikationsserver auf EJB-Basis zwar die Erstellung, Verteilung und Integration von E-Business-Java-Komponenten in der Middle Tier erleichtern, gilt dies nicht für die Daten. Um den Zielen der Markteinführung und Performance gerecht zu werden oder diese zu übertreffen, braucht man einen Middle-Tier Data-Cache Manager zur Ergänzung des EJB-Applikationsservers.

## Skalierung ohne Überlastung der Datenbank

Ein Middle-Tier Data-Cache Manager ist ein Produkt, das die Daten eines Unternehmens in einem persistenten Speicher zusammenführt und den Applikationsservern einen skalierbaren und verteilten Zugriff auf diese Daten ermöglicht. Ohne einen solchen Middle-Tier Data-Cache Manager kann man zwar eine B2B-Lösung (Business-to-Business) implementieren, hat jedoch keine Möglichkeit, diese zu erweitern, wenn die Anzahl der Benutzer oder die Anzahl der Transaktionen die Kapazität der Back-End-Systeme übersteigt. Die Möglichkeit einer dynamischen Weiterentwicklung des Systems wird dadurch in erheblichem Maße eingeschränkt.

Zu den B2B-Applikationen, die von der besseren Skalierbarkeit profitieren würden, die durch einen Middle-Tier Data-Cache Manager gewährleistet wird, gehören beispielsweise traditionelle Anbieter von Informationsdiensten, wie etwa Unternehmen für Versicherungsansprüche. Vor nicht allzu langer Zeit konnte nur der Service Provider selbst auf sein internes Information Warehouse zugreifen und dieses steuern. Heute jedoch fordern Kunden nicht nur den Online-Lesezugriff auf diese Daten, sondern auch die Möglichkeit, die Daten zu aktualisieren. In einer solchen Situation können die E-Business-Anwenderzahlen enorm hoch und schwer vorhersagbar sein, insbesondere wenn B2C-Anwendungen (Business-to-Consumer) am Informationsfluss beteiligt sind. In Abhängigkeit von verschiedenen Faktoren, wie z.B. Saison, Marktbedingungen usw., reicht die Anzahl der täglichen

Zugriffe von einigen tausend bis zu mehreren Millionen. Die gewöhnliche EJB-Server-Lösung würde darin bestehen, einfach weitere EJB-Server zu starten, um auf diese Weise neue Benutzer zu unterstützen. Ein solcher Notbehelf kann jedoch verheerende Auswirkungen auf den Datenzugriff haben. Durch mehrere hundert oder tausend zusätzliche Anwender werden die Datenbankverbindungen überlastet, wodurch sich die Systeme im gesamten Unternehmen verlangsamen können, wenn der Zugriff auf die Daten lediglich von Back-End-Systemen aus erfolgen kann.

Eine bessere Möglichkeit zur Bewältigung des Problems besteht darin, E-Business-Daten in einem Cache der Middle Tier zu speichern, wo der Zugriff in Speichergeschwindigkeit erfolgen kann. Durch diesen Ansatz wird der unvermeidbare Datenengpass vermieden, mit dem Systeme bei hohem Datenverkehr zu kämpfen haben, wenn sie einen einzigen Daten-Cache verwenden. Ein Middle-Tier Data-Cache Manager kann automatisch neue Daten-Caches hinzufügen, wenn man neue EJB-Server startet. Dies gewährleistet einen schnellen Datenzugriff, ohne die E-Business-Benutzer direkt mit den Datenquellen des Unternehmens verbinden zu müssen. Im Idealfall muss man sich dabei keine Sorgen um die Integrität der Daten machen, weil die Lösung über Funktionen verfügt, die eine transaktionsbezogene Konsistenz der Cache-Daten auch dann gewährleisten, wenn die Daten in mehreren Caches doppelt vorhanden sind. Daneben gibt es eine Reihe weiterer Leistungsmerkmale von verteilten Caches, die bei der Entscheidung für eine Middle-Tier-Lösung in Erwägung zu ziehen sind: Scheduling-Funktionen für Transaktionen, Weiterleitung von Anforderungen zwecks Lastverteilung sowie Bildung von Cache-Pools zur Minimierung der Ressourcennutzung. Im Vergleich zum Direktzugriff auf Unternehmensdatenbanken wird die Performance von E-Business-Systemen durch die Integration eines Data-Cache Managers in den EJB-Datenserver um ein Vielfaches gesteigert.

Das folgende hypothetische Beispiel, das auf einem tatsächlichen Modell mit realen Zahlen basiert, verdeutlicht die Arbeitsweise eines Middle-Tier Data Cache:

Die Systeme A und B schicken Informationen über Versicherungsansprüche an einen Web-Server. In einem Benchmark-Test verarbeiten die Systeme A und B 1.000, 3.000, 5.000 und 7.000 unterschiedliche Datenanforderungen. System A bestand aus einem Applikationsserver und einer relationalen Datenbank auf dem Back-End-System. Da eine relationale Datenbank verwendet wurde, war es erforderlich, Mapping-Code zu erzeugen. System B bestand aus einem Applikationsserver, einem Cache Manager und EJB-Speicher in der Middle Tier sowie einer relationalen Datenbank auf dem Back-End-System. Obwohl System B den gleichen Mapping-Code zwischen der relationalen Datenbank und dem Cache Manager erforderte, benötigte es

wegen seines Native-EJB-Speichers keinen Mapping-Code zwischen dem Applikationsserver und den Daten-Caches.

Bei einer niedrigen Anzahl von Transaktionsanforderungen war die Performance der beiden Systeme gleich gut. Als jedoch die Anzahl der Anforderungen stieg, lief System B doppelt so schnell wie System A. Am Ende lief System B mehr als 500 mal schneller als System A.

Wenngleich die Geschwindigkeit und Performance durch den Einsatz eines Data-Cache-Managers ohne Zweifel verbessert wird, wird diese Aufgabe nicht von allen Produkten gleich gut erledigt. Ein Data-Cache Manager, der über keinen EJB-Speicher verfügt, wird speichergebunden und muss die relationale Datenbank wiederholt aufrufen, um neue und manchmal redundante Daten anzufordern. Demgegenüber ist ein Data-Cache Manager mit einem EJB-Speicher in der Lage, gelesene und aktualisierte Daten zu speichern, sobald diese erstmals in verschiedene Caches geladen werden. Der Data-Cache Manager kann den Applikationsserver somit mit einer erstklassigen Performance versorgen, da für Daten mit Lesezugriff keine weiteren Daten mehr vom Back-End-System angefordert werden müssen. Der teure Mapping-Code wird nur bei Lese- und Schreibzugriffen mit Aktualisierung ausgeführt.

## Integration von Daten in der Middle Tier

Mit einem Data-Cache Manager wie beispielsweise Javlin™ von Object Design, lässt sich der Ausbau einer Applikation effizienter bewältigen. Um eine B2B-Applikation wie die hypothetische Anwendung für die Bearbeitung von Versicherungsansprüchen aufzubauen, werden zuerst die E-Business-Datenmodelle erstellt, indem Informationen aus dem gesamten Unternehmen zusammengetragen werden. Die Versicherungsapplikation führt Daten zu Tarifsätzen aus den vorhandenen Datenbanken mit Dokumenten zur Bearbeitung von Ansprüchen zusammen. Im nächsten Schritt werden neue E-Business-Daten wie z.B. Benutzerprofile oder Portale hinzugefügt. E-Business-Datenmodelle sind nie statisch. Deshalb muss das Datenmodell wahrscheinlich schon kurze Zeit nach der Implementierung der Applikation geändert werden, um z.B. neue Produkte hinzuzufügen oder ein neues Preismodell zu erstellen. Durch diese Änderungen muss möglicherweise eine zusätzliche und beschwerliche Aufgabe bewältigt werden: Datenmanagement - eine Aufgabe, die EJB-Applikationsserver nicht besonders schnell oder elegant erledigen können.

Wie kann man somit die Performance optimieren, ohne Einschränkungen bei der Flexibilität in Kauf zu nehmen? - Durch die Installation eines Middle-Tier Data-Cache Managers in jeden EJB-Server. Die Entwicklungszeit kann verkürzt werden, indem verschiedene Arten von Unternehmensdaten

automatisch in der Middle Tier konsolidiert werden, wo sie im EJB-bereiten Format gespeichert werden. Durch eine solche Lösung steht ein Platz zur Verfügung, an dem neue E-Business-Daten, wie z.B. Benutzerprofile, gespeichert werden, ohne vorhandene Unternehmensdatenbanken modifizieren zu müssen. Damit kann außerdem ein erweiterbares Daten-Framework bereitgestellt werden, das eine dynamische Hinzufügung von Datenkategorien und -attributen ermöglicht, ohne das System zu überlasten.

## Verwaltung mehrerer verteilter Caches

E-Business-Applikationen müssen in der Lage sein, mehrere hundert oder tausend Benutzer zur gleichen Zeit zu bedienen, wobei die Geschwindigkeit von entscheidender Bedeutung ist. Je schneller Transaktionen abgewickelt werden, desto schneller erhöht sich die Profitabilität. Mehrere Faktoren können Zeit- und Geldverluste für das System und das Unternehmen verursachen:

· Verbindung zu Back-End-Datenbanken
· Gleichzeitiger Zugriff auf Unternehmensdaten durch Legacy-Applikationen
· Wartezeit beim Lesen von Daten auf dem Plattenlaufwerk.

Durch Middle-Tier Data Caches werden Engpässe beim Zugriff auf die Datenbank beseitigt. Dabei können automatisch mehrere Daten-Caches erzeugt werden, aus denen die Enterprise Beans auf Daten zugreifen; angeforderte Daten, die sich in lokalen oder Back-End-Systemen befinden, werden automatisch in Caches gestellt, wo anschließend der Zugriff in Speichergeschwindigkeit erfolgen kann. Außerdem kann der Cache-Inhalt abgefragt, indiziert und sogar wiederhergestellt werden. Da diese Daten in einem sofort verwendbaren Format in der Nähe des Client gehalten werden, wird die Performance im gesamten Unternehmen gesteigert.

## Evaluation von Cache Managern

Bei der Evaluation von Middle-Tier-Cache-Systemen sollte man Systeme suchen, die sich zur Speicherung neuer Daten in einem Native-EJB-Format eignen. Dadurch entfällt die Erzeugung von Mapping-Code für die Daten des Unternehmens und die Daten des Kunden, der den Prozess überdies erheblich verlangsamen kann.

Bei nur einem Applikationsserver gewährleistet ein einziger Middle-Tier Cache eine exzellente Performance. Die Leistung kann jedoch beeinträchtigt werden, wenn die Applikationsserver durch zusätzliche CPUs erweitert werden. Zu berücksichtigen ist auch die Fragestellung, ob die Daten-Cache-Lösung mit der Skalierung der Applikationsserver mitwachsen kann. Durch die Möglichkeit, automatisch mehrere verteilte Caches bei Einhaltung der Konsistenz von Transaktionen zu erzeugen, wird

sichergestellt, dass die Daten-Caches flexibel genug sind, um die E-Business-Daten des Unternehmens zu modellieren. Außerdem sollte man die Möglichkeit haben, Daten geographisch, alphabetisch, nach Kategorien, nach Subsystemen, nach Lesezugriff, nach Aktualisierung oder anhand anderer individueller Systemanforderungen zu segmentieren.

Darüber hinaus sollte der Anbieter des Data-Cache Managers unabhängig vom Anbieter des EJB-Applikationsservers sein. Auf diese Weise kann man den besten Applikationsserver und die beste Daten-Cache-Lösung ermitteln. Die meisten Anbieter von Data-Cache Managern liefern eine erstklassige Performance und hochspezialisierte Produkte, sind jedoch gleichzeitig eng mit den Applikationsserver-Produkten integriert.

Das Aufkommen des Bereichs E-Business hat nicht nur zu einer Vielzahl neuer Herausforderungen geführt, sondern auch zum Entstehen einer ganzen Branche beigetragen, die Lösungen für diese Probleme anbietet. Datenserver wurden entwickelt, um die Interaktionen zwischen Unternehmen und ihren Kunden über das Internet zu erleichtern, und Data-Cache Manager wurden aus der Notwendigkeit geboren, die Leistung dieser Datenserver zu verbessern.

## Schlussfolgerung

Wenn man die Performance von Applikationsservern maximieren will, muss man sich wahrscheinlich mit dem Datenfluss in der Middle Tier befassen. Applikationsserver alleine können keine zuverlässige Performance liefern, wenn die Anzahl der Datentransaktionen einer Website steigt.

Mit einem Data-Cache Manager – der die Produktivität erhöht und die Komplikationen verringert - können viele Probleme, die durch die Anforderungen von E-Business verursacht werden, beseitigt werden.