

Vorsicht bei der Nutzung von Open-Source-Datenbanken!

Bei Big-Data-Projekten steht immer auch die Frage im Raum, ob mit Open-Source-Datenbanken gearbeitet werden soll oder nicht. Die Antwort hängt vom Service-Level ab, den es zu erfüllen gilt. Denn Datenqualität und Verlässlichkeit der Systeme haben hohe Priorität.



DER AUTOR

Adrian Hutzli
CEO, Intersys

Im Zuge der Verwendung neuester Technologien wie dem Internet der Dinge, 5G und anderen Innovationen steigen die Datenvolumina weiter massiv an, was Big-Data-Anwendungen wie Predictive Maintenance, künstliche Intelligenz oder Data Mining überhaupt erst möglich macht. Der Nutzen dieser Daten steht in der Diskussion in der Regel denn auch im Zentrum. Dabei gerät manchmal leicht in Vergessenheit, dass die beste Service-Idee rund um grosse Datenvolumen nur dann sinnvoll anwendbar ist, wenn das zur Verfügung stehende Datenmaterial auch wirklich korrekt, vertrauenswürdig und verlässlich ist.

In diesem Zusammenhang lohnt es sich, die Verwendung und die Art der Datenbanken für grosse Informationsmengen näher unter die Lupe zu nehmen. Denn die Datenbanksysteme sind, heute genauso wie früher, die Basis sämtlicher Datenverarbeitungsvorhaben. Hierfür stehen auch immer wieder Open-Source-Lösungen zur Diskussion. Denn immerhin gibt es mittlerweile ein Angebot an hochskalierbaren, mit kostenfreien Lizenzen erhältlichen Datenbanken, die eine Vielzahl an Funktionen zur Verfügung stellen, die einen professionellen Einsatz erlauben. So eignen sich etwa MongoDB, Elasticsearch, Apache Cassandra, MySQL oder Scylla bestens für diverse Anwendungszwecke und haben deshalb durchaus ihre Berechtigung. In den meisten Fällen garantieren regelmässige Updates, dass Fehler behoben und neue Funktionen eingeführt werden. Selbst für komplexe Fragen

steht in der Regel eine breite Community zur Verfügung, die bei Problemen unterstützend zur Seite steht.

Strenge SLA und Code-Weitergabe

Was aber, wenn wie bei vielen Big-Data-Vorhaben ein System gebaut werden soll, das für eine sehr hohe Verfügbarkeit ausgelegt sein soll und strenge Service Level Agreements erfüllen muss? Wie steht es dann um den Support und um allenfalls notwendige Korrekturen an der Software? Probleme können dabei bereits im Konstrukt von Open Source selbst stecken, etwa wenn ein System eine Schwachstelle aufweist. Verfügt nämlich derjenige, der das Leck gefunden hat, über kriminelle Energie, kann erheblicher Schaden entstehen. Dass die Schwachstelle behoben wird, ist zwar wahrscheinlich, aber nicht garantiert.

Bei der Notwendigkeit einer hochverfügbaren Anwendung muss also auch die Dauer zur Behebung eines Problems berücksichtigt werden. Deshalb sollte man sich wie bei der Verwendung einer Bezahl-Software gewahr sein, ob die Community den nötigen Support leisten kann und die Software die Aussicht hat, mehrere Jahre unterstützt zu werden. Andernfalls sollte man vielleicht besser eine proprietäre oder «Closed Source Code»-Software oder eine andere Open-Source-Lösung in Betracht ziehen. Ein weiterer Aspekt, den es zu berücksichtigen gilt, ist die Verpflichtung, Weiterentwicklungen an der Software an die Community zurückzugeben. Kommt man dieser Verpflichtung nach, besteht die Gefahr, geschäftskritische Informationen preiszugeben.

Fazit

Open-Source-Datenbanken sind für viele geschäftliche Anwendungszwecke durchaus geeignet und valable Alternativen zu proprietären Systemen. Muss aber eine Lösung sehr verlässlich sein, benötigt es auch eine entsprechende Absicherung. Bei kostenfreien Open-Source-Datenbanken kann eine mögliche Ergänzung zur Community die Zuhilfenahme eines Entwicklungsdienstleisters sein, der die interne IT mit Serviceverträgen unterstützt und so die nötigen Anforderungen garantiert und die benötigten Gewährleistungen übernehmen kann.



Bild: Dmitry Nikolaev / Fotolia.com