

Hochverfügbarkeit ist kein reines Infrastruktur-Thema

Bernd Usinger, GEBHARDT Sourcing Solutions AG

Kritische Geschäftsprozesse brauchen hochverfügbare IT-Systeme. Andernfalls können durch unvorhergesehene Ausfälle immense Schäden entstehen. Aber welche Hochverfügbarkeitsklasse ist die richtige, um einen digitalen Prozess optimal auszurichten? Und wie lässt sie sich erreichen? Intelligente Infrastruktur-Lösungen wie Oracle RAC sind effektive Hilfsmittel, um Unternehmens-Applikationen möglichst ausfallsicher zu gestalten, sie nähern sich dem Thema jedoch allein von der technischen Seite. Das greift schlicht zu kurz: Hochverfügbarkeit braucht ein ganzheitliches Denken.

Die Digitalisierung der Wirtschaft schreitet mit großen Schritten voran – man könnte fast meinen, sie trage Siebenmeilenstiefel: Kaum ein Geschäftsprozess kommt heute noch ohne steuernde IT-Systeme aus, die entscheidend zu seiner Wertschöpfung beitragen. Insbesondere in der Fertigungsindustrie entstehen durch die zunehmend IT-gestützte Vernetzung von Teilprozess-

sen, Maschinen oder ganzen Werken neue Umsatzpotenziale, die über die zukünftige Konkurrenzfähigkeit der hiesigen Industrie entscheiden.

Mit dem Grad der Digitalisierung eines Unternehmens steigt deshalb auch der Anspruch an die Verfügbarkeit seiner Systeme, weil das Schadenspotenzial bei Ausfall einzelner IT-Komponenten

inzwischen sprunghaft angestiegen ist. War beispielsweise der Ausfall eines E-Mail-Systems noch vor einigen Jahren auch über mehrere Tage hinweg gut zu verkraften, entsteht Unternehmen heute durch den Ausfall einer solchen IT-Anwendung oder eines ERP-Systems innerhalb weniger Stunden bereits ein riesiger Schaden – nicht selten in existenzgefähr-

dender Höhe. Insbesondere kritische Geschäftsprozesse erfordern deshalb einen verlässlichen IT-Betrieb, der vor riskanten Ausfällen sicher ist.

Das Stichwort lautet „Hochverfügbarkeit“. Es bezeichnet die Fähigkeit eines Systems, seinen Betrieb auch dann mit einer hohen Wahrscheinlichkeit aufrechtzuerhalten, wenn eine seiner Komponenten plötzlich ausfällt. Der Begriff findet heutzutage eine fast inflationäre Verwendung. Dabei ist Hochverfügbarkeit nicht gleich Hochverfügbarkeit: Eine Verfügbarkeit von 99 Prozent mag auf den ersten Blick hoch klingen, bedeutet aber für ein System, das rund um die Uhr im Einsatz ist, dass es letztlich pro Jahr immer noch stolze 87,7 Stunden lang ausfallen darf – das sind drei Tage und mehr als 15 Stunden.

Mit jedem Zehntelprozentpunkt steigt die Ausfallsicherheit um 8,76 Stunden pro Jahr. Von Hochverfügbarkeit spricht man allerdings erst ab einem Verfügbarkeitsniveau von 99,99 Prozent, womit die durchschnittliche Ausfallzeit nur noch rund 52 Minuten pro Jahr beträgt. Für die meisten Use Cases ist das ein tragbares Risiko – es gibt aber auch Anwendungsfälle, in denen selbst 52 Minuten Ausfall im Jahr nicht akzeptabel sind, etwa in der Notfallversorgung in Krankenhäusern, in der Luft- und Raumfahrt oder in der Kraftwerk-Steuerung. Hier muss eine Verfügbarkeit von 99,9999 Prozent gegeben sein, damit gewährleistet ist, dass die Funktion des Systems unter allen Umständen verfügbar ist. Aufgrund der vielen Nachkommastellen ist diese Hochverfügbarkeitsklasse im Englischen auch als „six nines“ bekannt. Die potenzielle Ausfallzeit lässt sich damit auf eine halbe Minute pro Jahr minimieren.

Je höher die Hochverfügbarkeitsklasse, desto höher die Investition

Dass Hochverfügbarkeit nicht ohne Investitionen zu erreichen ist, liegt auf der Hand. Dennoch sollten auch mittelständische Unternehmen sich davon nicht abschrecken lassen. Denn eine einzige Stunde Server-Ausfall kann im Zeitalter der Digitalisierung mitunter mehr kosten als eine komplette Hochverfügbarkeitslösung. Klar ist allerdings: Je höher die Hochverfügbarkeitsklasse, desto höher die Investitions-

und Betriebskosten für ein System. Aus diesem Grund kann es nicht das Ziel sein, jede einzelne Anwendung so verfügbar wie möglich auszuliegen. Es gilt vielmehr, die maximal tragbare Ausfallzeit gegen den maximal tragbaren Schaden beziehungsweise Datenverlust abzuwägen, um zu entscheiden, wie verfügbar das jeweilige System sinnvollerweise sein sollte.

Eine gute Entscheidungsgrundlage hierfür liefert die „Availability Environment Classification“ (AEC) der Harvard Research Group, eine Einteilung in sechs Hochverfügbarkeitsklassen: Die niedrigste Klasse, AEC-0 (Conventional), besagt, dass die Funktion eines Systems unterbrochen werden darf und die Datenintegrität nicht essenziell ist. Auch bei der Hochverfügbarkeitsklasse AEC-1 (Highly Reliable) kann die Funktion unterbrochen werden, die Datenintegrität muss jedoch gewährleistet sein. AEC-2 (High Availability) gibt an, dass die Funktion einer Anwendung nur innerhalb festgelegter Zeiten beziehungsweise zur Hauptbetriebszeit allenfalls minimal unterbrochen werden darf.

Noch weiter geht AEC-3 (Fault Resilient): Anwendungen in dieser Verfügbarkeitsklasse müssen innerhalb festgelegter Zeiten oder in der Hauptbetriebszeit ununterbrochen aufrechterhalten sein. Fehlertolerante Systeme dagegen (AEC-4/ Fault Tolerant) müssen ihre Funktion ununterbrochen halten, sodass ein 24/7-Betrieb gewährleistet ist. An höchster Stelle steht schließlich die Verfügbarkeitsklasse AEC-5 (Disaster Tolerant): Hier muss die Funktion unter allen Umständen verfügbar sein.

Wie aber findet ein Unternehmen nun heraus, welche Verfügbarkeitsklasse es im Einzelfall anstreben sollte? Um diese Frage sinnvoll beantworten zu können, ist an erster Stelle eine professionelle Analyse der betroffenen Geschäftsprozesse gefragt, wie sie ein fachkundiger IT-Dienstleister vornimmt: Welche von ihnen sind tatsächlich kritisch? Welchen Schaden erleidet das Unternehmen bei einem Ausfall der Anwendung? Oder anders gefragt: Wie lange darf die Anwendung im Störfall maximal ausfallen, damit der Schaden noch tragbar ist?

Wenn im Hochfrequenzhandel an der Börse mehrere Millionen Transaktionen pro Minute online durchgeführt werden, ist der Schaden schon bei einem zehnmütigen Ausfall des Systems immens und damit keineswegs tragbar. In ande-

ren Geschäftsfeldern dagegen, beispielsweise im Reporting, entstehen allenfalls Schäden in tolerierbarer Höhe, selbst wenn eine Anwendung einmal vier Stunden lang ausfällt. Hier für teures Geld in die höchstmögliche Verfügbarkeitsklasse zu investieren, steht schlicht nicht im richtigen Verhältnis.

Redundanz kritischer Systemkomponenten ist gefragt

Erst wenn ausgehend vom Geschäftsprozess eruiert wurde, wie ausfallsicher ein System tatsächlich sein muss, steht die Überlegung an, wie die entsprechende IT-Infrastruktur aufgebaut sein muss. In diesem Zusammenhang ist auch zu berücksichtigen, welchen Reifegrad die betroffene Anwendung hat: Soll sie gerade erst eingeführt werden oder ist sie bereits in der Optimierungsphase? Je höher der Reifegrad einer Anwendung ist, desto größer ist auch der Aufwand, um nachträglich eine höhere Verfügbarkeitsklasse zu erreichen.

Bei aller Stabilität und Qualität muss eine jede Unternehmens-IT deshalb zugleich ausreichend flexibel ausgelegt sein, um auf technische Neuerungen, veränderte Regularien oder die Entwicklungen des Marktes binnen kürzester Zeit reagieren zu können. Im Zeitalter der fortschreitenden Digitalisierung von Geschäftsprozessen stehen Unternehmen also vor der Herausforderung, einen Spagat zwischen der hohen Qualität und Verlässlichkeit ihrer IT-Systeme sowie deren Flexibilität schlagen zu müssen.

Hohe Verfügbarkeiten lassen sich im Allgemeinen erreichen, indem sogenannte „Single Point of Failure“-Risiken eliminiert werden. Dahinter verbergen sich diejenigen Systemkomponenten, deren Versagen den Ausfall des gesamten Systems nach sich ziehen würde. Soll eine Anwendung hochverfügbar ausgelegt sein, ist deshalb eine Redundanz aller kritischen Systemkomponenten anzustreben, um damit zugleich die Robustheit und Fehlertoleranz des Gesamtsystems zu erhöhen. Wie lassen sich die betroffenen Anwendungen also fehlertolerant innerhalb der geplanten Architektur abbilden? Wie lässt es sich erreichen, dass eine Applikation auch dann noch verfügbar ist, wenn ein Server oder ein Plattensystem im Storage ausfällt?

Eine verlässliche Lösung liefert das Oracle Real Application Cluster (RAC). Als zusätzliche Option der Oracle-Datenbank ermöglicht es Unternehmen, ihre Datenbanken ausfallsicher einzurichten, indem mehrere Knoten eines Rechnernetzes auf dieselbe Datenbank zugreifen. Im Falle eines Ausfalls einer dieser Knoten können sich Clients ohne Wiederanlaufzeit über einen der anderen Knoten mit dem System verbinden. Für den Anwender wiederum macht es keinen Unterschied, auf welchem der Knoten sein Zugang erfolgt.

Anders als bei klassischen Failover-Clustern wird die redundante Hardware bei der Oracle-Lösung nicht nur im Fehlerfall eingesetzt, sondern kann auch im Normalbetrieb genutzt werden: Alle Lasten können auf sämtliche Clusterknoten verteilt sein, sodass das RAC nicht nur für eine hohe Verfügbarkeit sorgt, sondern auch sehr flexibel skalierbar ist. Für Unternehmen entsteht dadurch mitunter ein deutlicher Performance-Gewinn.

Hochverfügbarkeit erfordert ein Ende-zu-Ende-Denken

Es ist allerdings zu kurz gedacht, das Thema „Hochverfügbarkeit“ allein von Seiten der Infrastruktur anzugehen. Denn es nützt in der Tat wenig, eine Datenbank mit dem Oracle-RAC von technischer Seite her hochverfügbar auszulegen, wenn die Kühlung des Server-Raums bei hohen Temperaturen nicht sichergestellt ist – oder wenn die zuständige IT-Belegschaft um 17 Uhr Feierabend macht. Hochverfügbarkeit betrifft Menschen und Systeme gleichermaßen und sie erstreckt sich über den gesamten Geschäftsprozess. Wer sich allein auf die technische Seite konzentriert, deckt damit nur einen Teilbereich ab und investiert teures Geld, ohne Systemausfälle zuverlässig verhindern zu können.

Stattdessen ist bei der Hochverfügbarkeit – wie so häufig in der Optimierung von Geschäftsprozessen – ein Ende-zu-Ende-Denken gefragt. Dieses lässt sich am besten von außen steuern, etwa durch einen versierten Dienstleistungspartner, der die internen Abläufe von einer externen Perspektive aus betrachten kann. In einem Auftakt-Workshop sollten der beratende Dienstleister und das betroffene Unternehmen zunächst eine klare Zielsetzung entwickeln, um welche Use Cases

es eigentlich geht, welche Anwendungen also überhaupt hochverfügbar ausgerichtet werden sollen. Dazu müssen zunächst einmal die Gefahrenpotenziale definiert werden, die einem Unternehmen durch eventuelle Systemausfälle entstehen können. Über ein fundiertes Risikomanagement lässt sich dann eruieren, welche IT-Abläufe damit innerhalb der Organisation im Einzelnen verbunden sind und auf welcher Rechenzentrums-Infrastruktur sie aufbauen, aber auch, welche räumlichen und personellen Ressourcen es braucht, um die Ausfallsicherheit so weit wie nötig zu minimieren und das System dabei dennoch so flexibel wie möglich zu halten.

Nicht immer ist es dabei sinnvoll, die eigene Infrastruktur kostenintensiv hochverfügbar auszurichten. Es kann unter Umständen besser sein, über eine verlängerte Werkbank im Sinne einer Co-Location oder gar einer Infrastruktur as a Service nachzudenken. Ein erfahrener Sourcing-Partner, der den Markt und seine Preise genau kennt, kann bei der „Make or Buy“-Entscheidung fundierte Hilfestellung leisten.

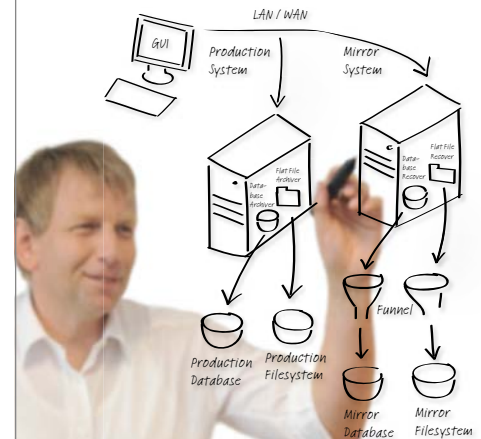
Fazit

Es ist deutlich absehbar, dass die Hochverfügbarkeit in den kommenden Jahren weiter an Brisanz gewinnt, denn im zunehmend vernetzten Industrie-4.0-Umfeld steigen die Anforderungen an die IT immer weiter. Hochverfügbare Systeme werden für Unternehmen gleich welcher Branche und Größe zum entscheidenden Wettbewerbsfaktor. Wer diese Dringlichkeit versteht und das Thema ganzheitlich denkt, kann mit den rasanten Entwicklungen der Digitalisierung auch in Zukunft Schritt halten.



Bernd Usinger
bernd.usinger@gebhardt-ag.de

Libelle BusinessShadow®



Unabhängig bezüglich

- Fehlerursache
- Entfernung
- Hardware / Architektur
- Komplexer Systeme

Schnelle Arbeitsaufnahme

- Mit konsistenten Daten
- Auf Knopfdruck
- Automatisiert
- ...

Hans-Joachim Krüger
Chief Technology Officer
Libelle AG

Erfahren Sie mehr:
www.Libelle.com/business



ORACLE Gold Partner



Libelle

Libelle AG
Gewerbestr. 42 • 70565 Stuttgart, Germany
T +49 711 / 78335-0 • F +49 711 / 78335-148
www.Libelle.com • sales@libelle.com