Mythen und Wirklichkeit bei Hochverfügbarkeit

Marco Metzlaff, Libelle AG

Je mehr Geschäftsapplikationen als „business-critical" oder „mission-critical" eingestuft sind, desto wichtiger ist es, deren Ausfallzeiten so gering wie möglich zu halten.

Applikationen basieren größtenteils auf relationalen Datenbanken. Zur Minimierung der ungeplanten Ausfallzeiten im Fall einer Katastrophe sind die Applikationen und Datenbanken gleichermaßen durch hard- oder softwarebasierende Lösungen zu schützen. Viele Versprechungen kursieren rund um das Thema „Disaster Recovery". Sie entpuppen sich bei genauerer Betrachtung als Mythen, weil sie größtenteils nur singuläre Punkte in der Technologie absichern.

Den Schutz geschäfts- oder unternehmenskritischer Applikationen gegen Katastrophen (englisch: Disaster Recovery, kurz DR) gehört zu den wichtigen, aber oft missverstandenen Aufgaben der Systembetreuer. Eine Implementierung von Infrastrukturmethoden – zum Beispiel mit Hardware, mit Snapshots oder mit einer Virtualisierungslösung – ist nicht in jedem Fall ausreichend, um komplexe Applikationslandschaften im DR-Fall zügig, konsistent und weitgehend verlustfrei wiederherzustellen.

Aufgrund von früheren, mittlerweile sicherlich überwundenen Schwächen in der Hardware legen viele Anwender den Fokus auf die infrastrukturelle Absicherung der jeweiligen Applikation. Dadurch entstanden im Markt aber auch einige Mythen über die DR-Fähigkeiten der IT. Die Realität sieht gerade mit einem Fokus auf die gesamte Applikationslandschaft anders aus als bei der Betrachtung spezifischer Hard- oder Software-Komponenten. Folgende Aussagen, die man oft in Diskussionen mit IT-Verantwortlichen findet, erweisen sich bei genauerem Hinsehen als Mythos:

1. Im Falle eines Disasters muss meine Applikation ganz ohne Datenverlust auf das Ersatzsystem umgeschaltet werden
2. Meine Systeme sollen sich ohne manuelle Eingriffe selbstständig umschalten
3. Ein Restore vom Band stellt nach einer Katastrophe alle meine Daten der Applikation vollständig wieder her
4. Disk-Snapshots können die Hochverfügbarkeit und Disaster-Recovery-Fähigkeit meiner Systeme sicherstellen
5. Eine Spiegelung der Datenbank mit Oracle DataGuard genügt für die Verfügbarkeit der Applikationen

Jede dieser Aussagen klingt recht plausibel, ist aber bei genauerer Betrachtung nur für einzelne Komponenten einer Applikations-Infrastruktur als Katastrophenvorsorge hilfreich. Dennoch sind sie bei vielen Anwendern wichtige Elemente des DR-Konzepts. Zudem muss darauf hingewiesen werden, dass zwar fast alle Anwender von Oracle-Datenbanken über ausreichende Backup- und Hochverfügbarkeitsstrategien verfügen, Konzepte zur Katastrophenvorsorge darin jedoch nicht immer im erforderlichen Umfang enthalten sind.

Die Konsistenz-Vorgabe

Klassisch bemisst sich die Qualität von Wiederanlauf-Szenarien sowohl mit den Größen der Recovery Point Objective (RPO), also wie viel Datenverlust ist im Fehlerfall tolerierbar, als auch mit der Recovery Time Objective (RTO), also wie lange dauert der Wiederanlauf. Diese Kriterien werden allerdings oft nur für einzelne Systeme definiert – müssten aber für die Wiederverfügbarkeit der Geschäftsprozesse festgelegt werden.

Die Verteilung der Daten und deren logischer Zusammenhang über eine Vielzahl von Systemen in Kombination mit einer ganzheitlichen Betrachtung der Daten über die Geschäftsprozesse stellt die IT vor die Herausforderung, Daten systemübergreifend konsistent wiederherstellbar zu machen.

Das Prinzip der „Logical Units of Work" (LUWs) funktioniert im Grunde nur innerhalb abgeschlossener Umfelder, etwa in der jeweiligen Datenbank. Schnittstellendaten und Daten in File-Systemen besitzen größtenteils keinen transaktionalen Konsistenz-Algorithmus. Die Erzeugung einer Gesamtsystemkonsistenz ist somit bei einer ganzheitlichen Betrachtung von Datenbanken, File-Systemen und Schnittstellen besonders schwierig. Konsistenz muss zudem quantifizierbar und überprüfbar hinterlegt und umgesetzt werden. Eine Angabe über RTO und RPO, die auf der Basis einzelner Systeme definiert und errechnet werden, ist aus Sicht der Geschäftsprozesse nicht ausreichend.

Für eine systemübergreifende Datenintegrität definiert die Recovery Consistency Objective (RCO) über alle Systeme hinweg die benötigten Konsistenzanforderungen. Im Detail beschreibt die RCO die erlaubte Abweichung wiederhergestellter Datenbestände nach einer Katastrophe. Sie gibt somit an, wie groß der Unterschied der Geschäftsdaten vor und nach dem De-

saster, verteilt über die beteiligten Systeme, sein darf und das sowohl qualitativ als auch quantitativ. Die RCO kann also Katastrophenschutzkonzepte auf Schwachstellen prüfen. Wird neben RTO und RPO auch die RCO in die Beurteilung der Katastrophenvorsorgemaßnahmen einbezogen, fällt es leicht, die eingangs genannten Mythen zu entlarven.

Der Mythos vom Umschalten einer Applikation ohne Datenverlust

Immer wieder erhalten Anwender das Versprechen, ihre IT-Umgebung würde durch die Entscheidung für eine bestimmte Hochverfügbarkeitslösung ohne jegliche Datenverluste wiederhergestellt werden können. Zwar können Hochverfügbarkeitslösungen für einzelne Server oder Applikationen auch einen gewissen Katastrophenschutz bieten, für eine vollständige IT-Umgebung können sie aber keine DR-Funktion erfüllen.

Im Katastrophenfall muss stets der Ausfall des kompletten Rechenzentrums betrachtet werden. Beim Starten der Ersatzsysteme müssen deshalb die wichtigen Geschäftsprozesse und die dafür erforderlichen Transaktionen in der IT-Umgebung im Mittelpunkt stehen. Dies zu erreichen ist wichtiger als der Verlust der Transaktionen während einer Wiederherstellungsprozedur.

Der Mythos vom bedienerfreien Umschalten

Manche Anwender erwarten von ihren Hochverfügbarkeits- und DR-Lösungen, dass sie den Störungsfall selbstständig erkennen und bei der Erfüllung bestimmter Kriterien automatisch auf die Ersatzsysteme umschalten. Ein automatisches Disaster Recovery ist technisch grundsätzlich realisierbar, entspricht jedoch nicht den Prinzipien des Business Continuity Managements (BCM). Im Rahmen von BCM-Prozessen muss der Katastrophenfall als solcher deklariert werden – sowohl hinsichtlich der Faktoren, die eine Katastrophe beschreiben, als auch hinsichtlich der Maßnahmen, die ergriffen werden müssen. Bei einer Katastrophe im Sinne der schweren Beschädigung des Rechenzentrums durch Überschwemmung, Brand oder ähnliches müssen stets auch manuelle Wiederherstellungsmaßnahmen definiert werden. Regelmäßige Disaster-Recovery-Tests – Erfahrungswerte bei Anwendern erlauben hier eine Empfehlung von zwei Tests pro Jahr – tragen dazu bei, dass die Wiederherstellung zügig verläuft und konsistente Datenstrukturen liefert.

Der Mythos Disaster Recovery mit Tape-Backups

Das klassische Backup auf Magnetband (Tape) ist nach wie vor eine wertvolle Maßnahme für den Schutz der Integrität von Daten, zur sicheren Speicherung und zur revisionssicheren Archivierung von Daten. Als Basisabsicherung gegen Katastrophen haben Magnetbänder jedoch entscheidende Nachteile: So lassen sich nach dem Wiederanlauf der Hardware bei ordnungsgemäß durchgeführten Backup-Läufen zwar die Restore-Läufe durchführen, für ein zügiges und vollständiges Restore einer komplexen Umgebung steht jedoch vielen Anwendern nicht die erforderliche Anzahl an Bandlaufwerken zur Verfügung. Die vollständige beziehungsweise für den Geschäftsbetrieb erforderliche Wiederherstellung dauert also zu lange. Des Weiteren bil-

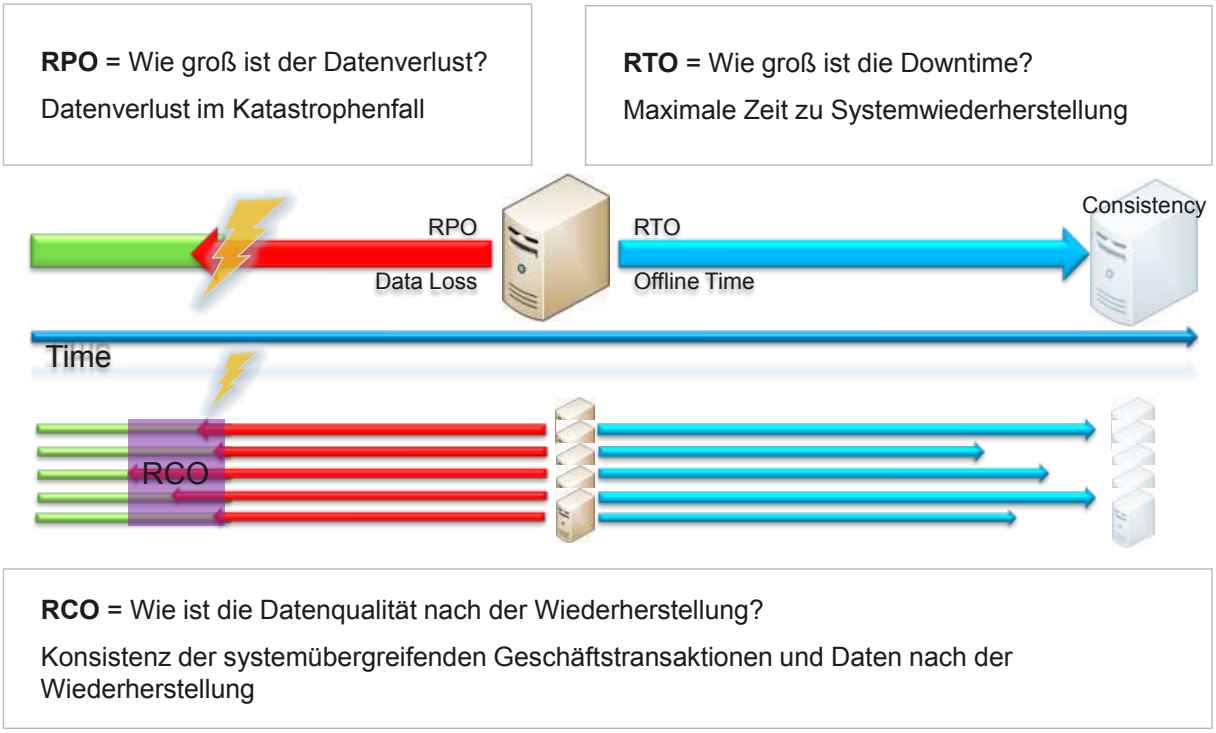


Abbildung 1: Recovery Point, Recovery Time, Recovery Consistency definieren den Datenverlust und die Zeit zur Systemwiederherstellung sowie die erforderliche Konsistenz der wiederhergestellten Applikationsumgebung

den Bänder oft nur den Zustand vor einigen Stunden ab, was einen relativ großen Datenverlust zur Folge haben kann. Die Bänder selbst sollten außerhalb des Schuttkegels der jeweiligen Gebäude aufbewahrt werden.

Das klassische Backup/Restore mit Bandtechnologien hat seine Berechtigung in lokalen IT-Landschaften geringer Komplexität, bei denen eine schnelle Wiederverfügbarkeit nicht die wichtigste Anforderung ist und die Anwender eine Wartezeit bis zur Wiederherstellung des letzten gültigen Datenstandes in Kauf nehmen können. Bereits bei vernetzten Systemen mit Zugriffen aus Filialen sind DR-Konzepte erforderlich, die über das Backup/Restore mit Magnetbändern hinausgehen.

Der Mythos von den Disk-Snapshots

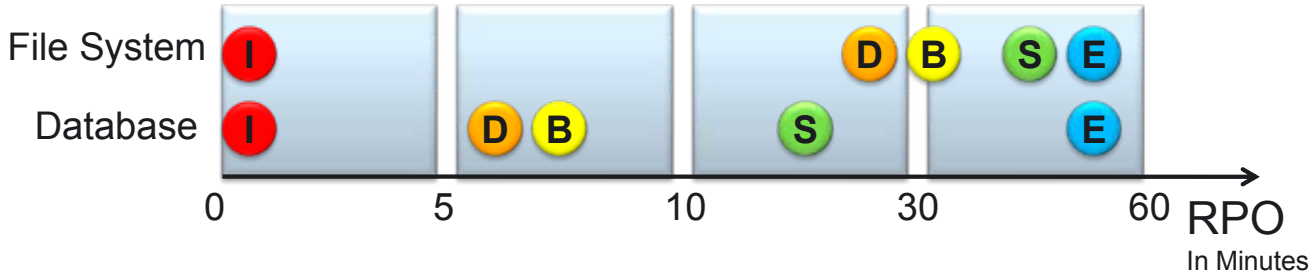
Anwender, und auch die Hersteller der jeweiligen Lösungen, sind vielfach der Meinung, die modernen Snapshot-Technologien wären gut geeignet, um die Daten schnell wiederherzustellen. Das ist grundsätzlich richtig, jedoch ist dabei zu beachten, dass Snapshots oft „space efficient“ gemacht werden und somit keinen vollständigen Datenbestand enthalten. Zudem werden Snapshots meist innerhalb ein und desselben Speichersystems angelegt, was die Absicherung gegen Katastrophen erschwert.

Für die Katastrophenvorsorge müssten die Snapshots auf ein Remote-Data-Center repliziert sein. Dafür sind wiederum die entsprechenden Anbindungen notwendig. Während Snapshots einzelne Dateien und vielleicht sogar größere Strukturen in einem Dateisystem wiederherstellen können, stoßen sie bei der Wiederherstellung von Datenbanken in einem DR-Fall zum Teil an die Grenzen. Die Datenbank muss ein Crash-Recovery durchführen, das mitunter auch mit einem Restore vom Band enden kann. Zum Teil müssen Offline-Logs der Datenbank mühsam aus unterschiedlichen Snapshots eingesammelt werden.

Zudem zeichnen Snapshots typischerweise nur die Veränderungen in den Daten auf, sodass zusätzliche „Initial Copies“ erforderlich sind, die im Wiederherstellungsprozess ebenfalls übertragen werden müssen. Lösungen, die für ein Backup-Szenario ausreichend und bei kleineren Störungen in der Hardware hilfreich sind, sind nicht notwendigerweise auch die richtige Lösung für die Katastrophenvorsorge.

Der Mythos von der DataGuard-Spiegelung

Maßnahmen zum Schutz gegen Bedien- und Softwarefehler basieren im Allgemeinen auf einer Spiegelung der Datenbank. Im Rahmen einer DR-Lösung ist eine Datenbankspiegelung jedoch nicht ausreichend, weil sie die Applikationsumgebung nicht berücksichtigen kann. Neben den Eskalationsprozessen zur Alarmierung von Mitarbeitern und zur Einleitung von Maßnahmen zum Katastrophenschutz müssen ja zusätzlich zur Datenbank auch die Applikationen, Einstellungsdaten, Schnittstellen und Dateisysteme gespiegelt und wiederhergestellt werden. Darüber hinaus stellt die Spiegelung der Oracle-Datenbank allein keinen ausreichenden Katastrophenschutz her, wenn auch Datenbanken anderer Hersteller im Einsatz sind, was in vielen gewachsenen Rechenzentren mit heterogenen IT-Landschaften der Fall ist.



- **System Classifications:**

Class	Description	Database	File System
E	End User	30-60 min	30-60 min
D	Vital Business Data	5-10 min	30 min
B	Business Support	5-10 min	30 min
I	Interfaces	~0	~0
S	Suporting Systems	10-30 min.	30-60 min.

Abbildung 2: Systeme werden im Katastrophenschutzkonzept entsprechend ihrer Rolle für die Aufrechterhaltung der Unternehmensaufgaben vorqualifiziert, wobei DR-Klassen (siehe Tabelle auf Seite 13) zur Anwendung kommen

Typical DR Classifications for Business Systems	
E: End User Connectivity Systems (Enterprise Portal) Keine vitalen Geschäftsdaten Nur User Properties Keine signifikante Änderung	RPO in Stunden RTO schnell RCO uninteressant
D: Vital Business Data Systems (ERP, CRM) Business Backbone, enthält Stamm- und Bewegungsdaten Signifikante Anzahl von Änderungen	RPO 5 - 30 min RTO schnell RCO signifikant für Geschäft und BI
B: Business Warehouse Systems (BI) Aggregate der Geschäftsdaten für den Entscheidungsprozess Massive Anzahl von Änderungen und Daten	RPO ~30 min, Delta Upload möglich RTO mittel RCO nicht insignifikant
I: Interface Systems (PI & Logistics) Permanente Änderung von elementaren Daten Hohe Abhängigkeit zu gekoppelten Systemen und Geschäftsprozessen	RPO ~0 RTO schnell RCO kann synchronisiert werden, oft aber nicht möglich
S: Supporting Systems (Solution Manager) Administrative Systeme, keine Geschäftsdaten	RPO in Stunden RTO langsam RCO meist insignifikant

Die entlarvten Mythen

Vieles, was heute als Disaster-Recovery-Lösung angeboten wird, ist bei genauerer Betrachtung nur für die Katastrophenvorsorge einzelner Infrastrukturkomponenten geeignet. DR-Konzepte funktionieren nur, wenn sie für den gesamten Verbund der Infrastrukturkomponenten und der Anwendungsumgebung Hardware- als auch Software basierende Werkzeuge und entsprechende Eskalationsprozeduren umfassen. DR-Konzepte für geschäftskritische Umgebungen müssen über die Wiederherstellung eines Systems hinausgehen und die Konsistenz der Daten und der Schnittstellen zwischen den beteiligten Systemen berücksichtigen. Nur so kann sichergestellt werden, dass mit der Wiederverfügbarkeit der Server auch die Applikationsumgebung vollständig und konsistent wiederhergestellt worden ist. Neben dem Recovery-Point-Objective und dem Recovery-Time-Objective sollte auch das Recovery-Consistency-Objective zur Beurteilung der DR-Konzepte herangezogen werden.

Kontakt:
Marco Metzlaff
marco.metzlaff@libelle.com

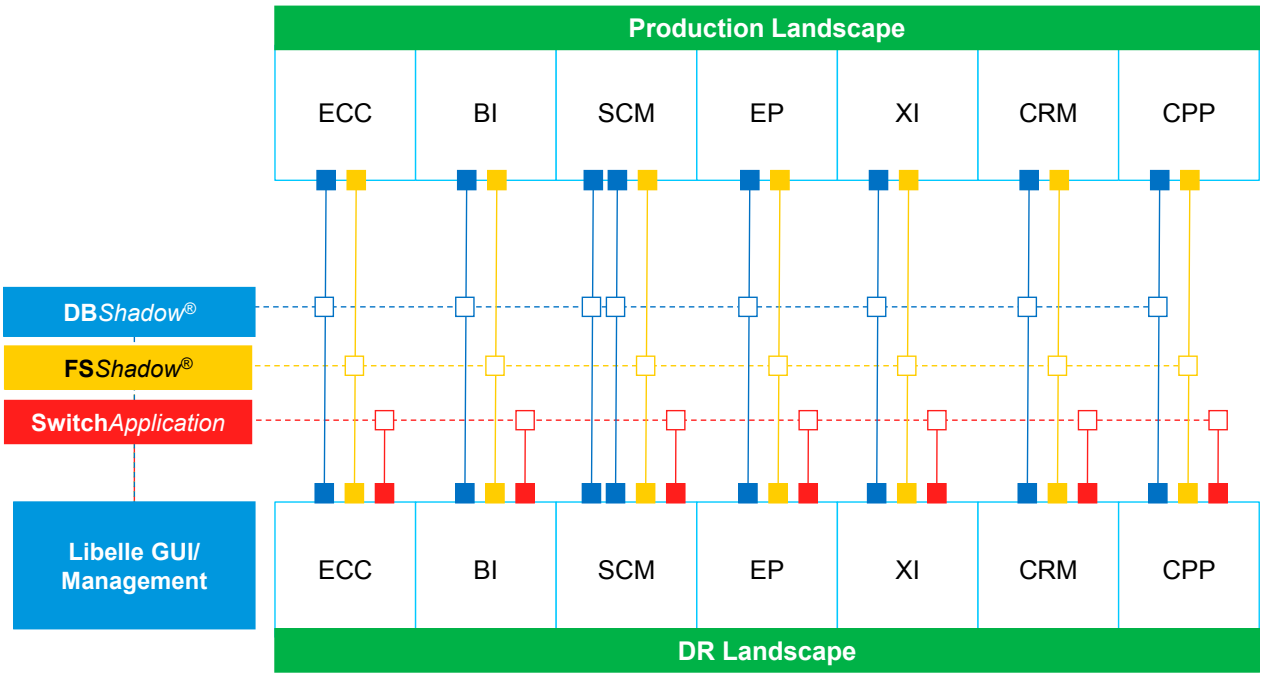


Abbildung 3: Komplexe SAP-Umgebungen benötigen viele Komponenten, die ein zeitlich synchronisiertes Disaster Recovery ermöglichen. Im Katastrophenfall steht so eine in die Prozesse integrierte Umschaltumgebung zur Verfügung