

manage *it*

[[IT - S t r a t e g i e n u n d L ö s u n g e n]]

Moderne ERP-Systeme

Komponentenbasierte Architektur
nach dem Baukastenprinzip

Business Intelligence neu erfunden

Paradigmenwechsel
beim Informationszugang

Auf den Mehrwert kommt es an

IT-Strategien der Zukunft

Drei Buchstaben mit neuer Relevanz

Marktentwicklung und Trends
bei DMS und ECM

Server-based Computing

Schlank und flexibel

Nutzen Sie unser Angebot für
Sonderdrucke oder E-Publishing-Dateien
von Artikeln dieser Ausgabe

Tel.: +49 8092 87543



Langzeitarchivierung

Records Management »auf Dauer«

Hersteller von Produktionsgütern und Dienstleistungen unterliegen nicht nur in der kaufmännischen Buchführung Dokumentationspflichten. In sicherheitskritischen Bereichen, wie der Kernkrafttechnologie oder chemischen Anlagen, müssen bestimmte Dokumente nach Regelungen von nationalen und internationalen Behörden über die Anlagenlebensdauer hinaus vorgehalten werden. Der Zeithorizont kann dabei 40 Jahre und länger betragen und überschreitet damit den Einsatzzeitraum der derzeit verfügbaren Hard- und Software.

Angesichts des langen Aufbewahrungszeitraumes ist die ungeschützte Speicherung von Dokumenten in großen Dateisystemen nicht sinnvoll, da die Datenbestände ohne geeignete Sicherheitsmechanismen unkontrolliert geändert werden können. Deshalb werden für die »dauerhafte« Speicherung von Daten häufig optische Medien in digitalen Archivelösungen eingesetzt. Bei einer Aufbewahrung von Dokumenten über die Lebensdauer der verwendeten Hard- und Softwarekomponenten hinaus spricht die Consultec von Archivierung »auf Dauer«.

Die Consultec hat mehrere Archivprojekte bei Herstellern und Betreibern komplexer technischer Systeme durchgeführt und dabei spezifische Erfahrungen gewonnen. So werden bei der Auslegung vielschichtiger technischer Systeme auch Entwicklungsaufgaben definiert, die die Kapazität und Performance der IT-Systeme bis in den Grenzbereich nutzen. Insbesondere bei Aufgaben, in denen große Gleichungssysteme (z.B. Finite Elemente) gelöst werden müssen, erreichen die Berechnungsergebnisse Dateigrößen, die in der Größenordnung derzeit erhältlicher Festplatten liegen und damit die Speicherkapazität optischer Medien übersteigen.

Neben den technischen Fragen gibt es das logische Problem, wie Doku-

mente nach so langen Zeiträumen erfolgreich recherchiert werden können. Administrative Merkmale wie Auftragsnummern, Organisationsbezeichnungen oder Kundennamen werden nach mehreren Jahrzehnten kaum geeignete Suchkriterien ergeben. Spannend ist natürlich auch die Frage, welche der heutigen Dateiformate in 40 oder 50 Jahren noch anzeigbar sein werden.

Im Kern müssen bei der Konzeption eines Langzeitarchivs deshalb drei konkrete Fragestellungen gelöst werden:

- II Wie kann eine zuverlässige Archivierung von Dokumenten »auf Dauer« erreicht werden?
- II Wie können Dokumente, deren Dateigröße das Fassungsvermögen der verwendeten Speichermedien übersteigt, archiviert werden?
- II Wie lässt sich sicherstellen, dass diese Dokumente über den langen Zeitraum auch recherchierbar und anzeigbar sind?

Archivierung »auf Dauer«. Die derzeit verfügbaren Hard- und Softwareprodukte (Datenträger, Archivsysteme und Datenbanken) lassen keine technologische Lebensdauer erwarten, die auch nur annähernd eine Archivierung »auf Dauer« ermöglicht. Im Laufe der Zeit müssen sowohl die eingesetzte

Hardware als auch die Software durch neuere Produkte ersetzt werden. Spekulationen über künftige technologische Entwicklungen liefern keine zuverlässige Basis für ein dauerhaftes Archivkonzept. Deshalb müssen im Unternehmen organisatorische Regelkreise installiert werden, die die Entwicklung von Technologie und Organisation beobachten und eine notwendige Migration von Hard- und Software veranlassen. Das betrifft eine Umspeicherung der archivierten Dokumente aufgrund der begrenzten Lebensdauer der Medien ebenso, wie den Wechsel der eingesetzten Hard- und Softwareprodukte und der Dateiformate. Da es keine endgültige Lösung des Archivproblems gibt, muss der Wandel organisiert werden.

Recherche und Anzeige. Ebenso wichtig wie die Technologie ist eine geeignete Ablagestruktur, die es etwa nach einem halben Jahrhundert gestattet, das Dokument erfolgreich zu recherchieren. Gemäß der Weisheit »weniger ist mehr« empfiehlt Consultec, ein einfaches Datenmodell zu entwickeln, das sich an einer relativ stabilen Begriffswelt orientiert.

Neben der Recherchierbarkeit kommt der Anzeigbarkeit der Dokumente große Bedeutung zu, da auch die

Dateiformate einer zeitlichen Änderung unterliegen. Anwendungsprogramme erzeugen in der Regel proprietäre Daten, das heißt Daten, die nur vom erzeugenden Programm selbst gelesen werden können. Zur Anzeige der Dokumente »auf Dauer« werden deshalb Dateiformate benötigt, die auf offenen Standards basieren und deshalb als langzeitstabil angesehen werden können.

Für die digitale Archivierung konventioneller Dokumente bevorzugt man meist darstellungstreue Dateiformate wie TIFF oder PDF, wie sie auch mit Scannern erzeugt werden können. Für Engineering-Daten, die bereits digital vorliegen, sind darstellungsneutrale ASCII-Dateien eine sinnvolle Alternative, da sie die digitale Bearbeitung der Daten erleichtern.

Diese abgeleiteten Dateien sind inhaltsgleich zur proprietären Ursprungsdatei und werden Langzeitrendition genannt.

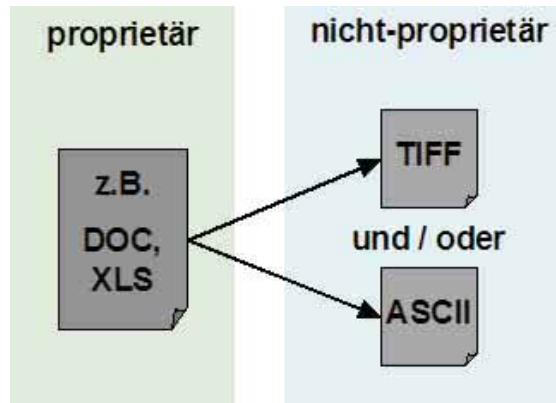
Archivierung großer und sehr großer Dateien.

Die derzeit gebräuchlichen optischen Speichermedien fassen unter 5 GByte pro Seite. Medien, die bis zu 30 GByte fassen, befinden sich in der Einführung. Gegebenenfalls existieren weitere Limitierungen in Hard- und Software. Da die zu archivierenden Dateien diese Größe übersteigen können, stellt sich die Frage, wie diese Grenze überwunden werden kann.

Am folgenden Beispiel soll die Archivierung großer Dokumente verdeutlicht werden: Die Datei – sei es nun eine Langzeitrendition in einem standardisierten Dateiformat oder eine proprietäre Anwendungsrendition – muss in Einzeldateien verarbeitbarer Größe zerlegt werden. Ein solches Verfahren wird Datei-Splitting genannt.

Die Ordnungsmäßigkeit der Auftrennung der Dateien während der Archivierung und später das Zusammenfügen der Teildateien zur Gesamtdatei werden durch die Bildung von HASH-Werten auf Gesamt- und Teildateien sichergestellt.

Ein wichtiger Konfigurationsparameter ist die Größe der Teildateien. Zu große Werte bewirken einen erheblichen »Verschnitt« bei der Ausnutzung



Bei der Langzeitrendition werden proprietäre Ursprungsdateien in inhaltsgleiche nicht-proprietäre Formate umgewandelt.

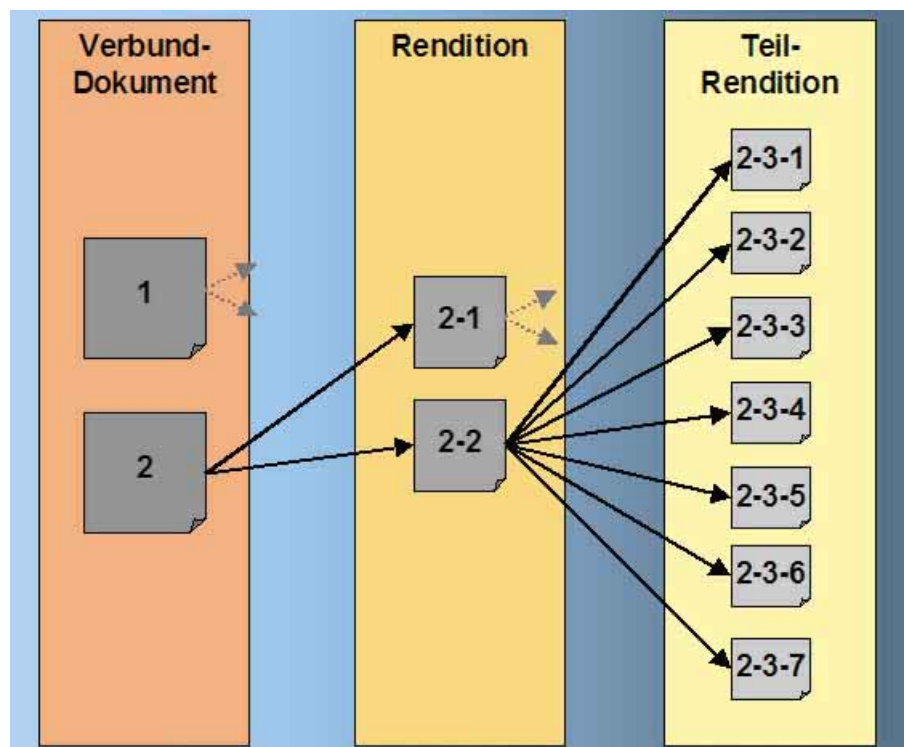
der Archivmedien. Zu kleine Werte ergeben eine Vielzahl von Teildateien, die die Performance des Archivsystems belasten. Deshalb sollte als Splittgröße nicht die maximal zu archivierende Dateigröße genommen werden, um einen optimalen Füllgrad der Speichermedien zu erreichen. In jedem Fall wird ein flexibler Mechanismus benötigt, der eine leichte Anpassung an zukünftige Mediengrößen ermöglicht.

Zur Erhöhung der Rechtssicherheit empfiehlt die Consultec eine Zertifizierung

der Archivelösung durch ein unabhängiges IT-Dienstleistungsunternehmen, wie beispielsweise den TÜV-IT. Die Zertifizierung setzt eine Verfahrensdokumentation auf Basis der PK-DML (Prüfkriterien für Dokumenten-Management-Lösungen) und ein erfolgreiches Audit vor Ort voraus.

Michael Hellmuth

Michael Hellmuth ist Senior Berater für die Unternehmensberatung Consultec Dr. Ernst GmbH



Große Dokumente werden archiviert, indem die Datei – sei es nun eine Langzeitrendition in einem standardisierten Dateiformat oder eine proprietäre Anwendungsrendition – in Einzeldateien verarbeitbarer Größe zerlegt werden. Ein solches Verfahren wird Datei-Splitting genannt.

3 Ausgaben

Einblick

Durchblick

Ausblick

f ü r d r e i z e h n f ü n f z i g !



[] **Ja**, ich bestelle drei Ausgaben » *manage it* « zum Preis von Euro 4,50 pro Ausgabe. Dieses Probeabonnement verlängert sich nicht automatisch.

Schicken Sie diesen Coupon an:

**ap Verlag GmbH
Postfach 1380
85554 Ebersberg**

oder faxen Sie die Seite einfach an die Nummer

+49 8092 87544

Titel: _____

Vorname: _____

Nachname: _____

Position: _____

Firma: _____

Straße: _____

PLZ: _____ Ort: _____

E-Mail: _____

Telefon: _____

Fax: _____