

Archivierung

Interview mit Paul Bathe, Business Technology Consultant, EMC Deutschland GmbH

Frage 1: Langzeitarchivierung scheint fast ausschließlich den großen Bandtechnologien vorbehalten. Können sich optische und magneto-optische Techniken sowie »exotische« Bandformate (VXA, Mammoth, Qic) noch behaupten?

Die Entwicklung der optischen Medien für die Archivierung ist stark beeinflusst von den HD-Formaten im Consumer-Bereich. Hier hat sich die Entwicklung auf Formatfragen, Zugriffsberechtigungen und Lizenzschutz der Inhalte konzentriert und nicht auf Langzeitstabilität und Kapazitätsentwicklung. Optische Medien können sich heute weniger behaupten als noch vor einigen Jahren. Da hatte noch fast jeder namhafte IT-Anbieter optische Archivlösungen im Portfolio, doch heute beschränkt sich die Zahl der Hersteller und Anbieter auf einige wenige.

Für optische Medien wie auch für exotische Bandformate gilt gleichermaßen: Sind nur wenige Anbieter im Markt wird sich das Produkt eher langsamer weiterentwickeln und damit ist ein Schritt halten mit den großen Disk- und Bandtechnologien sicher schwierig.

Frage 2: Compliance war lange Zeit Schlagwort und Triebfeder für Speicherprozesse und Archivierungsstrategien. Derzeit ist es ruhiger um dieses Thema geworden. Wie geht der IT-Anwender mit diesem Thema um? Ist er sensibilisiert genug oder vernachlässigt man das Thema?

Das Thema wird derzeit in den Märkten unterschiedlich behandelt. Während in sensiblen Bereichen wie Gesundheitswesen, Biotech, Chemie und Pharma als auch bei Banken und Versicherungen das Thema aufgrund klar definierter Prozesse und Vorschriften im Fokus steht, verzögert sich in klassischen Industriebereichen das Thema Compliance. Der Grund hierfür liegt in der fehlenden Defini-

“Sicher fängt eine gute Archivierungsstrategie mit der Definition von konzernweit einheitlichen Regeln und Prozessen an.“

Paul Bathe, Business Technology Consultant, EMC Deutschland GmbH

tion von konzernweit einheitlichen Compliance-Regeln. Diese Aufgabe stellt für die Konzerne eine Mega-Aufgabe dar, die sich nur mittelfristig umsetzen läßt. Kleinere Unternehmen sind da oft schon weiter und hier können wir durchaus wachsende Nachfrage für Archivspeichersysteme mit entsprechender „Compliance“-Zertifizierung beobachten. Das Thema Compliance hat bei den Unternehmen nach wie vor hohe Bedeutung, nur bei der Einführung und Umsetzung wäre eine konsequentere und schnellere Umsetzung wünschenswert.

Frage 3: Wie sieht eine gute Archivierungsstrategie aus? (Hierarchisch, welches Medium, Migrationszyklen)

Sicher fängt eine gute Archivierungsstrategie mit der Definition von konzernweit einheitlichen Regeln und Prozessen an. Hier wird dann schnell klar, daß man mit den bereits bestehenden verteilten Archiven auf unterschiedlichen Plattformen die gesteckten Ziele kaum erreichen wird. Im zweiten Schritt sollte daher eine Konsolidierung der Anwendungen und Archivplattformen erfolgen. Bisher war die Zusammenführung der verschiedenen Systeme im Filesystem bzw. HSM eine oft im Markt vorzufindende Lösung. Doch hat sich diese Lösung wohl eher aufgrund von Kostendruck als aufgrund einer guten Archivierungsstrategie entwickelt. HSM bietet in der Regel keine Indexierung, wenig Sicherheitsfunktionen und lässt kaum eine auf Basis von Prozessen und Workflows gesteuerte Archivierung zu, sodass sich hier in Zukunft zentrale ECM-Systeme mit zentral angeschlossener Archivplattform als die bessere Lösung erweisen werden.

Die Auslagerung von Daten auf Medien ist für eine gute Archivierungsstrategie kaum zu empfehlen. Offline-Medien sind schwierig zu verwalten, speziell wenn nach Ablauf der Daten eine zertifizierte Löschung vorgenommen werden soll bzw. Einzelobjekte gelöscht werden müssen. Sie sind auch kaum zu kontrollieren, wenn sie über Standorte hinweg ausgelagert und transportiert werden. Eine Archivierungslösung bestehend aus einem zentralen gesicherten Repository mit direkt angeschlossener Archivplattform, das auf Basis von Prozessen und Workflows die Regeln umsetzt, stellt sicher die bessere Strategie dar.

Eine gute Archivierungsstrategie betrachtet ebenso den Aufwand für Migrationzyklen, die bei einer Langzeitarchivierung von zehn und mehr Jahren unabdingbar sind. Verteilte Archive und Medien sind hier kaum wirtschaftlich, zumal in diesem Fall noch wesentlich kürze Zyklen für Daten-Refresh betrachtet werden müssen. Systeme mit zentral angeschlossener Archivierungsplattform und Tools zur automatisierten Datenmigration und Technologie-Refresh sind daher zu bevorzugen.

Frage 4: Was sollte der IT-Suchende für die Auswahl der Technik/Strategie beachten?

Folgende Fragen sind aus unserer Sicht für die Auswahl der Technik und Strategie zu beachten.

Wie wird sichergestellt, dass sich der Informationsgehalt der zu ablegenden Daten auf dem Archivierungssystem nicht ändert? Wichtig ist hier eine Funktionalität, die über das Prüfen von Hardware-Fehlern hinausgeht. Das System muss die Daten permanent auf Integrität prüfen. Dabei muss die Information während der gesamten Lebenszeit überprüft und ggf. selbstständig aus der Kopie wieder hergestellt werden.

Können für verschiedene Applikationen jeweils unterschiedliche Aufbewahrungsfristen verwaltet werden? Wie werden die unterschiedlichen Aufbewahrungsfristen systemtechnisch umgesetzt?

Welche Möglichkeiten gibt es, um auf Änderungen der Rahmenbedingungen schnell und effektiv reagieren zu können (z.B. beim Verkürzen der gesetzlichen Aufbewahrungsfristen etc.)? Betrachten Sie dabei die Wirkung auf neue als auch auf bereits archivierte Dokumente!

Was passiert nach Ablauf der Aufbewahrungsfrist mit einem Dokument im Speichersystem?

Für ein Enterprise Archive ist es wichtig zu den zu archivierenden Dokumenten Metadaten mitzugeben, um später gegebenenfalls schnell und effektiv über alle Applikationen hinweg nach Informationen zu suchen. Inwieweit unterstützt Ihr Produkt diese Funktionalität?

Bietet das Produkt die Möglichkeit zum Accounting an?

Frage 5: Können Disk-Systeme wirklich eine Alternative zu anderen Archivtechniken sein? Gibt es hier wirkliche WORM-Funktion?

Ja, absolut. Wir haben 2003 als erster Hersteller ein revisionssicheres Plattenarchiv herausgebracht, das von Wirtschaftsprüfern zertifiziert wurde und heute das meistverkaufte plattenbasierte Archivsystem weltweit ist: die EMC Centera. Weitere Systeme im Markt haben das seitdem bestätigt. Die WORM-Funktionalität wird bei Disk-basierten Systemen über Software bzw. Firmware geregelt. Andere Systeme setzten HASH-Verfahren ein, um die Unveränderbarkeit der Daten verlustfrei über mathematische Verfahren sicherzustellen. Diese Systeme bieten zudem den Vorteil weiterer Compliance-Funktionen wie „Litigation Hold“ und „Event Base Retention“. Mit „Litigation Hold“ werden vom Prüfer ausgewählte Dokumente für den Zeitraum einer Untersuchung gegen das Löschen der Daten Verfahrensgeschützt. Mit der Funktion „Event Base Retention“ wird die finale Aufbewahrungsfrist erst mit Eintreffen eines bestimmten Ereignisses vergeben.

Frage 6: Gewinnt LTO nun sämtliche S-DLT-Nutzer? Gibt es andere Optionen auf der Suche nach Ablöseformaten?

EMC hat sich schon immer klar für Disk-basierte Speichersysteme eingesetzt. Nun stellen wir fest, dass die Kunden zunehmend die Vorteile von Plattenspeichersystemen gegenüber Tape-basierten Systemen im Bereich Langzeitarchivierung und dem Backup von Daten erkennen. Mit der Einführung der TB-Platten, intelligenten Verfahren zur Deduplizierung sowie der Verfügbarkeit stromsparender Prozessoren kann das Festplattensystem gegenüber dem Bandsystem auch in wirtschaftlicher Hinsicht punkten. Weitere Vorteile sind beispielsweise die Einzelobjekt-Ablage und wesentlich höhere Zugriffsgeschwindigkeit auf einzelne Objekte und Datenblöcke.

Im Compliance-Umfeld werden auch mit LTO-Laufwerken Refresh-Zyklen notwendig, die teilweise eine größere Auslegung der Systeme notwendig machen, als es für das normale Archivaufkommen zunächst notwendig wäre. Die Grenze der Wirtschaftlichkeit verschiebt sich daher auch hin zu großen Kapazitäten zugunsten der Disk-basierten Archivsystemen, die derzeit noch Band-basierten Systemen vorbehalten waren.

Frage 7: Verwaltung, Integrität und Lesbarkeit stellen oft Probleme in der Langzeitdatenvorhaltung dar. Wie lassen sich hier Schwierigkeiten vermeiden?

Für die Verwaltung und Integrität haben sich HASH-Code Verfahren und sogenannte CAS-Systeme (Content Addressed Storage) bereits bestens bewährt. Die Lesbarkeit ist eine Frage des Formats und es ist zu hoffen, dass sich auch hier einheitliche Standards entwickeln. In der Praxis haben sich Bildformate wie TIFF und JPG, als auch PDF/A im Bereich Belegarchivierung etabliert. Solange aber keine verbindlichen Standards vorliegen, läßt sich die Lesbarkeit und Integrität der Daten für Langzeitsicherheit nur durch das Archivieren des Originaldokuments, d.h. die Ablage des Dokuments in dem Format, in dem es von der Anwendung erzeugt wurde und parallel als unveränderbarer Archivbeleg als TIF/JPG oder PDF/A sicherstellen. Die einzelnen Versionen der Anwendungssoftware und Betriebssysteme sind dann ebenso als Image im Archiv vorzuhalten. Gängige Praxis ist hier z.B. die Generierung und Ablage der Images über VMware.

Frage 8: Gibt es derzeit sich abzeichnende Trends in der Archivierung?

Trend 1: ISO Images bzw. Container-Schreibweisen werden zwar noch bei einem Archivaufkommen von vielen kleinen Objekten und Datenblöcken verwendet, wenn das aus Performance-Gründen oder aber aufgrund von nicht beherrschbarer Anzahl von Objekten notwendig ist. Die Schreibweise über ISO Images weicht aber immer mehr der Einzelobjektanlage, insbesondere weil die Daten Life Cycle Verwaltung von Einzelobjekten wesentlich komfortabler und wirtschaftlicher ist und Archivsysteme im Markt verfügbar sind, die in den Bereich von mehreren Milliarden Objekten skalieren.

Trend 2: Direkte Archivanbindung aus den führenden Applikationen heraus wie z.B. die Verbindung über die Archive Link SAP Schnittstelle an die Archivplattform.

Trend 3: Die Verwaltung eines zentralen Index für die konzernweite Suche über alle Anwendungen, Bereich und Standorte. Dieses ist auch für die Umsetzung von Compliance Regeln eine oft gestellte Forderung.

Trend 4: Ein weiterer Trend stellt die Standardisierung der Schnittstelle für Archivplattformen über XAM dar. Unser Unternehmen hat diese Initiative bereits in 2004 gestartet und. XAM wird sicher den Archivmarkt stark beeinflussen. Alle namhaften Hersteller von Archivsystemen haben sich bereits zu XAM committed. EMC bieten bereits heute die XAM Version 1.0 mit den CAS-Produkt an. Es ist zu hoffen, dass neben den Archivplattformen mehr und mehr Anwendungen diesen Standard unterstützen werden. Der Austausch der Daten, die Datenmigration und weitere Prozesse können dann für die Langzeitarchivierung wesentlich vereinfacht und automatisiert werden.



EMC Deutschland GmbH
Am Kronberger Hang 2a
65824 Schwalbach/Taunus
06196 / 4728 0
www.emc2.de

Machen Sie den nächsten Schritt:

Revisions-sichere digitale Langzeitarchivierung nimmt in der IT zunehmend eine wichtigere Rolle ein. Ein umfassendes Konsolidierungs- und Speicherkonzept ist dabei von unschätzbarem Wert. Informieren Sie sich, wie EMC Ihnen dabei helfen kann.