

## **KOST.07**

### **Kolloquium „Archivtaugliche Speicherinfrastruktur“**

#### **Thema: Blackbox**

#### **Inhalt**

1	Einleitung .....	2
2	Die Referate .....	2
2.1	Ilka Stahlberg, Jörg Homberg, Brandenburgisches Landeshauptarchiv .....	2
2.2	Raphael Mettan, Christian Kübler, IBM Schweiz .....	4
3	Diskussion .....	4
3.1	Wie werden die archivischen Anforderungen erfüllt? .....	4
3.2	Welches sind die wichtigsten technischen Aspekte? .....	5
3.3	Mit welcher Grössenordnung ist für die Kosten zu rechnen? .....	5
3.4	Realisierbarkeit: Ist das etwas für die Staatsarchive? .....	6

## 1 Einleitung

Die dritte Veranstaltung des KOST-Kolloquiums "Archivtaugliche Speicherinfrastruktur" befasste sich mit dem Thema "Blackbox".

Unter einer Blackbox verstehen wir eine Speicherlösung, bestehend aus Hard- und Software, deren interner Aufbau für den Benutzer im technischen Sinn transparent, also nicht einsehbar ist. Der Benutzer kommuniziert mit der Blackbox über eine klar definierte Schnittstelle. Die Blackbox sorgt selber für den Erhalt des archivierten Bitstroms; dieser kann deshalb für die Lebensdauer des Systems als gegeben angenommen werden. Der Benutzer muss sich nur über die Migration zu einem neuen System zu gegebener Zeit Gedanken machen.

Das Thema wurde im KOST-Kolloquium aus verschiedenen Blickwinkeln beleuchtet: aus Sicht der Anwender sowie der Anbieter, mit Blick auf zwei verschiedene Systeme. Zunächst präsentierten Ilka Stahlberg, Abteilungsleiterin, und Jörg Homberg, wissenschaftlicher Dokumentar, die Erfahrungen des Brandenburgischen Landeshauptarchivs mit seiner Lösung für die digitale Archivierung, die für die Speicherung auf dem System Centera von EMC beruht. Danach stellten Raphael Mettan und Christian Kübler eine Blackbox-Speicherlösung von IBM vor, die DR 550.

## 2 Die Referate

### 2.1 Ilka Stahlberg, Jörg Homberg, Brandenburgisches Landeshauptarchiv

Die digitale Archivierung im Brandenburgischen Landeshauptarchiv (BLHA) konzentriert sich zur Zeit auf die Archivierung aus Dokumentmanagementsystemen (DMS). Die konzeptuelle Basis dazu bilden das DOMEA-Konzept und seine Erweiterungsmodule<sup>1</sup>. Konkrete Arbeiten seit 2003 umfassten einen Test zum Aktenaustausch zwischen verschiedenen DMS mittels der XDOMEA-Schnittstelle<sup>2</sup>, die Mitarbeit an Geschäftsordnung und Registraturrichtlinien sowie vor allem die Entwicklung einer produktunabhängigen XML-Schnittstelle für die Archivierung aus DMS. Durch die Aufnahme dieser Aktivitäten in den Masterplan eGovernment der Landesregierung Brandenburg wurden sie im Akten-Lifecycle der Verwaltung verankert.

Um das Ziel einer reversionssicheren Archivierung zu erreichen, müssen verschiedene Anforderungen erfüllt sein. Dazu gehören neben Überlegungen zu Formaten auch solche zur Speicherinfrastruktur. Der Entscheid fiel für das System Centera der Firma EMC.

---

<sup>1</sup> [http://www.kbst.bund.de/cln\\_028/nn\\_836960/Content/Standards/Domea\\_Konzept/domea\\_no\\_de.html](http://www.kbst.bund.de/cln_028/nn_836960/Content/Standards/Domea_Konzept/domea_no_de.html)

<sup>2</sup> <http://www.koopa.de/produkte/XDOMEA/index.html>. Die XDOMEA-Schnittstelle wird bis im Frühjahr 2008 überarbeitet; dabei soll die Archivierungsschnittstelle XARCHIV integriert werden.

Die Centera ist ein festplattenbasiertes Speichersystem<sup>3</sup>. Es erscheint dem Anwender gegenüber als Blackbox; das bedeutet, es gibt eine einzige, standardisierte Zugriffsmöglichkeit auf die Centera über ein API (Application Programming Interface). Die interne Organisation dieser Speicherbox ist gegenüber dem Benutzer transparent, er braucht sich nicht darum zu kümmern. Bei der Speicherung sendet ein Anwendungsserver ein Objekt (Binary Large Object, BLOB) über dieses API an die Centera, welche darauf eine eindeutige Content-Adresse berechnet und diese an die Anwendung zurückgibt. Die Content-Adresse beruht auf einem Hashwert des Objekts, dient also zugleich als Integritätsprüfung. Beschreibende Daten werden in einem Content Description File (CDF) abgelegt, das ebenfalls eine Content-Adresse erhält. Erst nach der Generierung von Spiegeldateien werden die nun als c-Clip bezeichneten Objekte (BLOB + CDF) der Langzeitarchivierung zugeführt. Die Konsistenz der Daten beruht auf der RAIN-Technologie (Redundant Array of Independent Nodes) und wird von der Centera ständig automatisch überwacht. Anzumerken ist dazu, dass dies zu einem relativ hohen Stromverbrauch führt. So musste im Serverraum eine zusätzliche Klimatisierung installiert werden, da das System die Raumtemperatur stark ansteigen liess.

Das BLHA führt eine Reihe von Gründen an, die zum Entscheid für die EMC Centera als Speichersystem geführt haben. Von grosser Bedeutung ist die automatische Konsistenzüberprüfung. Damit ist die archivische Anforderung der Unveränderbarkeit der Daten in weitgehender Weise erfüllt. Im Oktober 2007 bestätigte eine Havarie die Krisensicherheit der Centera: Bei der automatischen Überprüfung entdeckte das System spät abends einen schwerwiegenden Fehler und meldete diesen sowohl dem Administrator als auch EMC. Am folgenden Morgen diagnostizierte ein Supporter einen Stromversorgungsschaden in einem Knoten. Tags darauf traf das Ersatzteil ein und wurde eingesetzt. Die Datenkonsistenz war nie in Gefahr; nach dem Einsetzen wurden die Daten vollständig repliziert und das Speichersystem war wieder voll funktionsfähig.

Für die Lösung spricht zudem der geringe Administrationsaufwand. Da sich die Blackbox Centera selber überprüft, beschränkt sich der Aufwand für den menschlichen Administrator auf die Lektüre eines Kontrollmails pro Tag. Im Havariefall ist mehr Aufwand erforderlich, der sich jedoch ebenfalls in Grenzen hält.

Dies trägt auch zu einer speziellen Verteilung der Speicherkosten bei: Die Centera verursacht einmalig grosse Investitionskosten, danach jedoch minimale Administrationskosten und vor allem keinen zusätzlichen Personalaufwand. Frau Stahlberg wies darauf hin, dass es im Bereich der digitalen Archivierung in der Regel schwierig sei, neue wiederkehrende Ausgaben ins Budget zu setzen, dass es dagegen einfacher sei, einmalige grössere Ausgaben zu rechtfertigen. Diesen Budgetierungsvorgaben entspricht eine Blackboxlösung sehr gut. Das BLHA hat für die aktuelle Lösung (8 TB Speicherplatz installiert, erweiterbar bis 28 TB, inkl Servicevertrag für 3 Jahre) 106'000 Euro bezahlt. Allerdings ist es zur Zeit nicht möglich zu sagen, wie lange diese Lösung den Bedürfnissen genügen wird, d.h. wann eine neue Blackbox beschafft und die Daten migriert werden müssen. Eine erste Migration von einem Testsystem auf ein produktives System wurde jedoch im Frühling 2007 bereits erfolgreich und problemlos durchgeführt.

---

<sup>3</sup> Es besteht auch die Möglichkeit, eine Tape-Library einzubinden, was im BLHA nicht getan wurde.

Die aktuellen Perspektiven zielen in erster Linie auf die Beschaffung einer zweiten Centera zur Spiegelung der archivierten Daten. Von EMC ist angekündigt, dass mit grösseren Speichervolumen zu deutlich günstigeren Preisen und mit massiv gesenktem Stromverbrauch gerechnet werden kann.

## **2.2 Raphael Mettan, Christian Kübler, IBM Schweiz**

IBM ist im Gebiet der digitalen Langzeitarchivierung verschiedentlich engagiert, unter anderem durch die Mitarbeit in Forschungsprojekten, als Anbieter von Speichersoftware sowie Speicherhardware. Vorgestellt wurde die Infrastrukturlösung IBM DR 550. Wie die zuvor vorgestellte Centera ist auch die DR 550 eine Blackbox, die über das Tivoli Storage Manager (TSM) API, angesprochen wird. Intern verwaltet ein Controller ein SAN, an welches zur Zeit bis zu 168 TB Diskspace angeschlossen werden kann. Besonderen Wert legt IBM auf die Möglichkeit, einen Tiered Storage<sup>4</sup> zu verwalten, d.h. die archivierten Objekte je nach den konkreten Anforderungen auf verschiedenen Medien abzulegen. Dazu können neben den in der DR 550 vorhandenen Disks auch Tape-Libraries eingebunden werden. Der Vorteil einer (teilweisen) Verwendung von Tape als Speichermedium liegt vor allem in den massiv kleineren Kosten, verursacht unter anderem durch geringeren Stromverbrauch und günstigere Wartung. Die Policy-Verwaltung ermöglicht es, verschiedene Policy-Domänen für unterschiedliche Arten von Daten zu definieren, die dann unter anderem auf verschiedene Medien gespeichert werden.

Die Funktionsweise der DR 550 ist der oben beschriebenen der Centera grundsätzlich sehr ähnlich. Archivobjekte werden über das API dem Speicher übergeben; anders als die Centera verwendet die DR 550 nicht Content Addressing, sondern verwendet einen Schlüssel als Identifikator.

## **3 Diskussion**

### **3.1 Wie werden die archivischen Anforderungen erfüllt?**

Die Diskussion konzentrierte sich zunächst auf den Gegensatz zwischen Disk und Tape und darauf, inwiefern die beiden Medien wesentliche archivische Anforderungen erfüllen. Das BLHA begründete seinen Entscheid für eine reine Disk-Lösung mit der erhöhten Integritätssicherheit, die durch die automatische Konsistenzüberprüfung erreicht wird. Diese ist auf Tape so nicht möglich. Die Administration einer Tape Library lässt sich zwar ebenfalls automatisieren. So kann die Lesbarkeit der Bänder mit sogenannten Audits sichergestellt werden. Die Überprüfung der Integrität einzelner Dateien, wie bei Centera, findet aber nicht statt, müsste also vom *Application Layer* übernommen werden. Das ist aber eigentlich überflüssig, da beim Lesen der Tapes bereits blockweise mit Checksummen geprüft wird, dass die Daten unverändert sind. Beide Systeme müssten aber eigentlich doppelt angelegt und die Daten repliziert werden, da in beiden Fällen Worst-Case-Szenarien denkbar sind. Anzuführen ist ferner, dass der erwartete Speicherplatzbedarf eines KOST-Mitgliedsarchivs in den nächsten

---

<sup>4</sup> Hierarchische Speicherverwaltung. So erlaubt es z.B. Tivoli® Storage Manager, die Daten nach unterschiedlichen Zugriffs- und Sicherheitskategorien auf unterschiedlichen Geräten und Medien zu halten. Damit wird jeweils die für diese Datenkategorie geeignetste und günstigste Speicherung eingesetzt.

Jahren noch eine Zehnerpotenz unter den kleinsten Tape-Library-Lösungen liegt (siehe dazu unten).

IBM brachte als Argumente für Tape einerseits die Kosten vor, andererseits die Erkenntnis (die im Rahmen dieses Kolloquiums auch bereits geäussert worden war), dass staatliche Archive die hohe Verfügbarkeit einer Online-Speicherlösung gar nicht brauchen.

### 3.2 Welches sind die wichtigsten technischen Aspekte?

Die beiden diskutierten Lösungen, Centera von EMC und IBMs DR 550 beruhen auf den gleichen Prinzipien und Funktionsweisen. Die Unterschiede sind punktuell: Centera benutzt Content Addressing<sup>5</sup>, DR 550 einen nicht-hashbasierten Schlüssel; IBM legt grosses Gewicht auf hierarchische Speicherung mit der Möglichkeit der Auslagerung gewisser Bestände auf Tape, EMC setzt hauptsächlich auf Disk.

Die wesentlichen Merkmale des Speicherkonzepts "Blackbox" sind aus beiden Präsentationen klar hervorgekommen. Das Speichersystem tritt den archivischen Anwendungen gegenüber völlig transparent auf. Angesprochen wird es über ein definiertes API. Dieses kann ein Objekt entgegennehmen, einen Identifikationsschlüssel ausgeben und bei Vorweisen dieses Schlüssels das Objekt wieder ausgeben. Die Details der Speicherung kümmern den Anwender nicht. Diskbasierte Systeme kontrollieren die Datenintegrität dauernd und selbständig.

Die Referenten betonten, dass eine offene Schnittstelle zwischen Applikation und Speicherplattform (z.B. XAM, 'eXtensible Access Method' oder 'X-Access Method') wünschenswert wäre. Heute ist diese Schnittstelle proprietär, bei EMC das *Centera-API* bzw. bei IBM das *TSM-API*. In der Regel wird auch ein NFS<sup>6</sup>/CIFS<sup>7</sup>-Protokoll zur Verfügung gestellt.

### 3.3 Mit welcher Grössenordnung ist für die Kosten zu rechnen?

Das BLHA hat, wie oben erwähnt, knapp über 100'000 EUR bezahlt für ein auf Centera basierendes Komplettsystem mit 7 TB Speicherplatz inklusive Wartungsvertrag für drei Jahre. Daraus Kosten pro TB und Jahr zu errechnen, fällt nicht leicht, weil die Laufzeit dieser Lösung noch offen ist.

Christian Kübler gab als ungefähre Grössenordnung für ein kleines, tape-basiertes DR 550-System an, dass für die DR 550 inkl. Software und für eine Tape-Library im Umfang von 40 TB mit je etwa 35'000 CHF zu rechnen sei. Für die Ausfall- und Katastrophensicherung müssten zwei solche Systeme beschafft werden, was demnach Kosten im Bereich von 140'000 CHF verursachen würde.

---

<sup>5</sup> Content Adressed Storage (CAS) ist ein spezielles Speicherverfahren, das einen direkten Zugriff auf einzelne Objekte ermöglicht und gleichzeitig die Unveränderbarkeit der gespeicherten Information sicherstellt. Um eine Zuordnung zwischen Inhalt und Speicherplatz machen zu können, kommen Hash-Funktionen zum Einsatz.

<sup>6</sup> Network File System (abgekürzt NFS) ist ein von Sun Microsystems entwickeltes Protokoll, das den Zugriff auf Dateien über ein Netzwerk ermöglicht.

<sup>7</sup> Common Internet File System (kurz CIFS) wurde 1996 von Microsoft eingeführt. CIFS baut auf NetBIOS over TCP/IP (kurz NBT) und SMB auf.

Charakteristisch ist, dass die Beschaffung einer Blackbox relativ hohe einmalige Investitionskosten, dafür während der Lebensdauer des Systems sehr tiefe Administrationskosten verursacht.

### **3.4 Realisierbarkeit: Ist das etwas für die Staatsarchive?**

In diesem Zusammenhang wurden zwei Themen berührt. Zunächst fiel auf, dass die Grösse des Archivs (d.h. der Umfang der zu archivierenden Daten) eine entscheidende Rolle spielt bei der Wahl des geeigneten Systems. Die KOST schätzt, dass ein mittleres Schweizer Staatsarchiv in den nächsten zehn Jahren einen Speicherbedarf im einstelligen TB-Bereich haben wird. Für solche Grössenordnungen sind die erwähnten Tape-basierten Lösungen zu gross, währenddem einfache Diskssysteme – einfachere als die an dieser Veranstaltung vorgestellten – darauf besser zugeschnitten sind. Falls sich jedoch mehrere Archive zur Realisierung einer Speicherlösung zusammenschliessen, erreichen sie Datengrössen, die die Verwendung von Tape nahe legen würden, sofern die übrigen damit verbundenen Implikationen ebenfalls gewollt sind.

Angesprochen wurde schliesslich die Frage, ob die Archive einer Blackbox vertrauen können, oder ob nicht vielmehr die Anforderung besteht, über die Funktionsweise der Speicherinfrastruktur im Detail informiert zu sein. Frau Stahlberg anerkannte, dass es vor zehn Jahren wohl noch unmöglich gewesen wäre, eine solche Lösung in einem Archiv durchzusetzen. Inzwischen haben jedoch alle, auch privat, Erfahrungen mit laufenden Systemen gemacht; es erscheint nun auch dem Laien plausibel, dass ein laufendes System eine hohe Sicherheit darstellt.