

Sauver les CD

Spezifikation zum Projektpunkt "Speicherbox im Archiv"

1 Zusammenfassung

2 Ausgangslage

Teil des Projekts *Sauver les CD* ist die Unterstützung der Archive bei der Sicherung der Dateninhalte von ephemeren Datenträgern auf einer geeigneten Speicherplattform. In vielen Archiven ist es über die eigenen Informatikdienstleister nicht möglich oder zu teuer, geeigneten Speicherplatz für diesen Zweck zu gewinnen. Die in etlichen Archiven bereits praktizierte Lösung, CDs und DVDs auf mobile USB-Festplatten zu kopieren, ist eher als heikel zu betrachten: nicht so sehr wegen möglicher Plattenausfälle als wegen der Gefahr eines unbeabsichtigten Überschreibens oder Löschens beim Manipulieren der Platten. In diesem Zusammenhang wurde die im parallelen Projekt *arcun*¹ für die *Bitstream Preservation* verfolgte Lösung, Mietspeicherplatz von mehreren Archiven gemeinsam zu nutzen, als valable Alternative diskutiert.

An der KOST-Aufsichtskommissionssitzung vom 24. September wurde das Projekt *arcun* jedoch vorläufig sistiert. Genau genommen soll vertiefend abgeklärt werden, ob anstelle einer reinen *Bitstream Preservation* eine Gesamtdienstleistung "Digitale Archivierung" für die beteiligten Archive nicht die vorteilhaftere Lösung sein könnte. Zu diesem Zweck sollen den Mitgliedarchiven Vor- und Nachteile beider Lösungen erläutert werden. Nach einem Richtungsentscheid im Frühjahr soll die eine oder andere Lösung realisiert werden. Auch bei einer optimistischen Einschätzung ist mit einer Umsetzung frühestens Mitte 2010 zu rechnen.

Aus Sicht der KOST gibt es zwei Überbrückungslösungen für die Speicherung:

1. Die Daten werden auf einem **File Server** der kantonalen Informatik abgelegt. Da diese Server grundsätzlich der Büroautomatisation dienen, ist der Speicherplatz dort in der Regel teuer (30'000 TB/Jahr²) und nur begrenzt verfügbar (100 GB max.). Diese Lösung bietet sich also vor allem für Archive mit bescheidenen Datenmengen an.
2. Bei grösseren Datenmengen drängt sich der Kauf einer **Speicherbox** (bzw. zweier Speicherboxen) mit einer Kapazität bis etwa 4 TB auf. Die Kosten

¹ Projektspezifikation *arcun*: Ziel und Zweck des Projektes *arcun* (rätoromanisch für "Speicher" oder *archival repository for collaborative use*) ist es, den Mitgliedarchiven der KOST (18 kantonale Staatarchive, das Schweizerische Bundesarchiv und das Landesarchiv Liechtenstein) preisgünstigen Speicherplatz für die digitale Langzeitarchivierung (LZA) zur Verfügung zu stellen. Die KOST übernimmt dabei Spezifikation, *Auditing* und Inkasso für die gewählte Lösung. Der Datenverkehr zwischen den beteiligten Archiven und dem Speicherplatzanbieter bzw. dem Speicherservice erfolgt direkt.

² 30'000 TB/Jahr und 100 GB: beide Angaben Informationen aus den Gesprächen bei der Bestandesaufnahme in den Archiven.

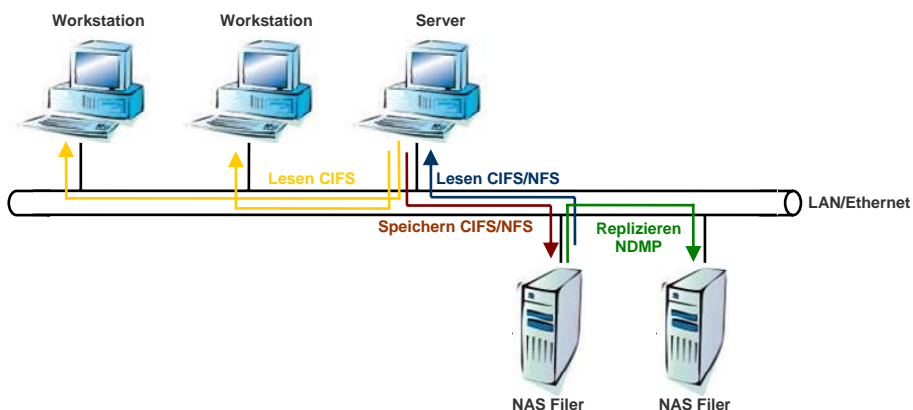
sollten sich dabei für zwei Speicherboxen und Software auf nicht mehr als 10'000 sfr.³ belaufen.

3 Speicherbox im Archiv

Für die zweite Lösung bietet die KOST Unterstützung bei Evaluation, Kauf und Installation an. Die KOST möchte einem Pilot-Archiv bei der Installation einer solchen Lösung zur Seite stehen. Die Lösung könnte dann von den andern interessierten Archiven begutachtet und übernommen werden. Der Vorteil dabei ist, dass nicht jedes Archiv selber eine Evaluation vornehmen muss und dass durch ein gemeinsames *Reviewing* der gewählten Lösung die Archivtauglichkeit zu einem hohen Grad gewährleistet ist.

3.1 Konzept

Die Lösung soll aus zwei NAS Speicherboxen⁴ mit je 3-4 TB Speicherplatz bestehen. Die beiden Boxen werden ins Intranet des Archivs integriert. Die Zugriffsberechtigung soll den Zugriff auf den Speicherplatz auf einen (oder ausgewählte) Rechner im Archiv begrenzen. Von diesem Rechner aus soll das Speichern und Lesen der Daten vorgenommen werden. Die eingesetzte *Storage Management Software* soll den Inhalt der einen Box automatisch auf die zweite Box spiegeln.



Aufbau der Speicherboxlösung

3.2 Anforderungskatalog

Aus diesem Grundkonzept erarbeitet die KOST-Geschäftsstelle einen Anforderungskatalog für die Lösung "Speicherbox". Die KOST wird dazu die Erfahrungen der Archive einbeziehen, die bereits mit einer ähnlichen Lösung vertraut sind. Der Anforderungskatalog wird folgende Themen umfassen:

³ Im Projekt **arcun** wurde für *full service bitstream preservation* ein Preis von 4500.- sfr./TB/Jahr ermittelt. Hardwarekosten sind nach einem verlässlichen Schätzwert etwa 1/4 der Gesamtkosten. Bei drei TeraByte und drei Jahren Laufzeit kommen wir auf Hardwarekosten von etwa 10'000.- sfr.

⁴ *Network-attached storage (NAS) is file-level computer data storage connected to a computer network providing data access to heterogeneous network clients.*
siehe: http://en.wikipedia.org/wiki/Network-attached_storage

- Betriebssystem
- NAS Software
- RAID-Controler/RAID-Level
- Festplatten-Typ
- Netzwerkanbindung
- Benutzeroberfläche

Der Anforderungskatalog wird im Zirkulationsverfahren den Archiven zur Ergänzung und Kommentierung zugänglich gemacht.

3.3 *Evaluation*

Die KOST evaluiert aufgrund des Grundkonzepts und des gemeinsam erstellten Anforderungskataloges eine mögliche Hardware/Software-Lösung. Im Workshop "*hands-on*" vom **15. Dezember 08** werden die besten Lösungen den interessierten Archiven vorgestellt. Die KOST wird eine Bewertung der Lösungen nach transparenten und nachvollziehbaren Kriterien vornehmen, sodass dann ein Entscheid gefällt werden kann.

3.4 *Installation*

Bei einem der an der Speicherbox-Lösung interessierten Archive wird mit Unterstützung der KOST die vorgeschlagene Lösung installiert und getestet. Das beteiligte Archiv muss dazu die Basisinfrastruktur, Strom und Netzwerkananschluss bereitstellen und einen Zugriffsrechner bezeichnen. Je nach Situation im Archiv wird die Lösung zuerst bei der KOST installiert und getestet und erst anschliessend an das Archiv ausgeliefert.

Nach der Installation soll eine *Review*- und Testphase auch den andern Archiven erlauben, die gefundene Lösung zu verifizieren.

3.5 *Kosten und Aufwand*

Die finanziellen Aufwendungen sollen für beide Boxen und Software den Betrag von 10'000 sFr nicht überschreiten.

Personeller Aufwand entsteht dem Archiv durch Bereitstellen der Infrastruktur, durch Unterstützen der KOST bei der Installation und durch das archivinterne *Reviewing*. Im Pilot-Archiv ist dafür mit etwa 6-8 PT zu rechnen. Im Gegenzug erhält das entsprechende Archiv die Unterstützung der KOST.

4 **Weiteres Vorgehen**

Dieses Papier soll den beteiligten Archiven als Grundlage für Gespräche mit ihren jeweiligen Informatikdienstleistern dienen und sie unter den gegebenen Umständen zu einem Mitmachen (auch als Pilot-Archiv) animieren.

Wir schlagen vor, dass die Archive das weitere Vorgehen bis zum Werkstattbericht am **17. November 08** in ihrem Hause besprechen und dass bis zum 15. Dezember 08 geklärt ist, welches Archiv die Rolle des Pilot-Archivs übernehmen möchte. Die KOST wird vor dem 17. November 08 eine entsprechende Rückfrage bei den Archiven durchführen.