

KOST.07

Kolloquium „Archivtaugliche Speicherinfrastruktur“

Thema: Speicherplatz mieten

Inhalt

1	Zur Einleitung	2
2	Die Referate	2
2.1	Manfred Abplanalp, Risk Manager; Luzius von Salis, Sales and Marketing von COLT Telecom AG.....	2
2.2	Hannes Heckel, Senior Product Manager von FAST LTA AG.....	2
3	Diskussion	4
3.1	Vertraglich-organisatorische Aspekte: Wie können die Leistungen abgesichert werden? Welche Arten von <i>Service Level Agreements</i> (SLA) sind möglich?	4
3.2	Zertifizierung: Wie stehen die Anbieter zur Diskussion über Trusted Digital Repositories und entsprechenden Zertifizierungsbemühungen?.....	4
3.3	Technische Fragen: Wie erfolgt der Zugriff auf die Daten?	5
3.4	Kosten: Mit welcher Grössenordnung ist zu rechnen?.....	5
3.5	Realisierbarkeit: Ist das überhaupt etwas für die Staatsarchive?	5

1 Zur Einleitung

Die erste thematische Veranstaltung des KOST-Kolloquiums "Archivtaugliche Speicherinfrastruktur" befasste sich mit dem Konzept "Speicherplatz mieten". Dabei wird der benötigte Speicherplatz, abgesichert durch langfristige Verträge, bei einem staatlichen oder privaten Anbieter gemietet. Der Anbieter übernimmt alle notwendigen Vorkehrungen für die Erhaltung des Bitstromes. In der gesamten Kolloquiumreihe fokussieren wir auf diesen Teil des *Archival Storage*-Prozesses. Die Aufgaben der *Digital Archiving Software*, etwa die Bewirtschaftung, Erhaltung, Migration etc. sind nicht Gegenstand der Untersuchungen (siehe dazu die Zusammenfassung der Einleitungsveranstaltung).

Zwei Firmen wurden eingeladen, in kurzen Referaten ihr Angebot vorzustellen: COLT Telecom AG¹ und FAST Long Term Archiving².

2 Die Referate

2.1 Manfred Abplanalp, Risk Manager; Luzius von Salis, Marketing & Business Service Director von COLT Telecom AG

Die Firma COLT ist ein Leistungserbringer und verkauft keine eigenen Produkte. Sie bietet eine Grundinfrastruktur für Netzwerk- und Datenservices an. Dazu unterhält sie ein eigenes, europaweites Glasfasernetz und eine Serie von Hochsicherheits-Datencentern. Hauptsächliches Zielpublikum ist die Finanzwelt (Banken, Versicherungen etc.); entsprechend ist die Grundinfrastruktur ausgestattet (unabhängige Stromversorgung, Sicherheit, Redundanz etc.).

COLT geht davon aus, dass es keine Standardlösungen für die Datenarchivierung gibt. Deshalb analysiert die Firma eingehend die Risikolandschaft mit ihren potentiellen Kunden. Untersucht werden dabei unter anderem die Risikofaktoren höhere Gewalt, menschliche Fehlhandlungen, technisches Versagen, organisatorische Mängel, Vorsatz und Systeme. Auf Grund dieser Analyse wird danach eine Lösung erarbeitet und angeboten.

Gegenüber potentiellen Kunden argumentiert COLT mit ihrer jahrelangen spezialisierten Erfahrung und mit dem Vertrauen wichtiger Kunden aus der Banken- und Versicherungswelt. Erfahrung mit Langzeitarchivierung im Sinn einer unbeschränkten Aufbewahrungsdauer hat die Firma bisher nicht; für eine Bank sowie für die eigenen Geschäftsdaten sind aber mittelfristige Archivierungslösungen implementiert.

2.2 Hannes Heckel, Senior Product Manager von FAST LTA AG

Im Gegensatz zu COLT ist FAST kein Datencenterbetreiber, sondern Spezialist und Produktanbieter für die Langzeitarchivierung, worauf sich die ursprünglich anderweitig tätige Firma seit Anfang 2006 spezialisiert. Sie bietet Online-Datensicherung als

¹ <http://www.colt.net/>

² <http://www.fast.ag/>

Alternative zu anderen Technologien an. Dazu hat sie einen Disk-Controller entwickelt, der WORM³-Funktionalität auf herkömmlichen SATA-Festplatten⁴ implementiert. Durch einen Kodierungs-Algorithmus wird mit doppelter Datenmenge eine neunfache Redundanz erreicht. Durch *Content Addressed Storage*⁵ werden Duplikate erkannt. Zudem ist im Controller ein *Digital Audit* integriert: Die Integrität der archivierten Daten wird automatisch alle 30 Tage geprüft; bei Fehlern wird der ganze betroffene Disk abgeschaltet und aus den redundanten Daten wieder rekonstruiert. Den Datenbesitzer braucht dies nicht zu kümmern; er kann aber im Rahmen des Reporting detaillierte Angaben zu Problemen und Ausfällen erhalten.

Mit dieser Technologie will FAST eine äusserst geringe Wahrscheinlichkeit für den Kompletterverlust der archivierten Daten, niedrige und vorhersehbare Kosten und wenig Administrativaufwand erreichen. Für die Dimensionen, welche die KOST für ein mittleres Staatsarchiv als realistisch erwartet (10 TB Daten, jährlicher Zuwachs 10%), rechnet FAST mit Kosten in der Grössenordnung von 0,15€ pro GB und Monat. Dabei muss beachtet werden, dass diese Berechnungen auf einem Sammelsystem beruhen. Für Staatsarchive ist unter Umständen ein dediziertes System erforderlich, d.h. Infrastruktur, auf der nur Archivdaten lagern.

Befragt zur Roadmap von FAST gab H. Heckel an, dass sie mit Ihrer FAST-Technologie mit der Entwicklung der Festplattentechnologie Schritt halten wollen. Da die gesamte Lösung nur auf einem speziellen Disk-Controller beruht und kein eigenes Dateisystem benutzt, kann für die Zukunft auch Abwärtskompatibilität erwartet werden.

³ "Write Once, Read Many times": Speichermedien, die nur einmal beschrieben werden können; gespeicherte Daten können weder abgeändert, überschrieben noch gelöscht werden. Siehe auch <http://de.wikipedia.org/wiki/WORM>.

⁴ http://de.wikipedia.org/wiki/Serial_ATA.

⁵ "Mit dem Content-Addressed-Storage-System wird auf gespeicherte Information nicht über ihre Lage auf dem physikalischen Medium, sondern über den Inhalt der Information zugegriffen." (http://de.wikipedia.org/wiki/Content-Addressed_Storage)

3 Diskussion

3.1 Vertraglich-organisatorische Aspekte: Wie können die Leistungen abgesichert werden? Welche Arten von Service Level Agreements (SLA) sind möglich?

COLT unterstreicht, dass die Lebensdauer des Anbieters ein hohes Risiko darstellt, gegen das sich ein Kunde absichern muss. Empfohlen wird, das Konkursrisiko über Retentionsrecht abzusichern: Damit übergehen die Rechtstitel an den Systemen im Konkursfall an den Kunden. Für COLT ist das Eigentum der Infrastruktur ein wichtiger Bestandteil der langfristigen Sicherheit. FAST weist auf die Vervielfachung der mit der Lebensdauer verbundenen Probleme hin, die durch Auslagerung von Bereichen und Prozessen an Dritte entstehen. Bei Fusionen ist die Rechtsnachfolge zwar geklärt, aber die Konsequenzen sind dennoch unklar.

Der Zeithorizont für die eigene Infrastruktur liegt für COLT bei etwa 20 Jahren; darüber hinaus zu planen ist nicht sinnvoll. Der Zeithorizont für Verträge mit Kunden liegt in der Regel bei fünf Jahren. Es sind verschiedene Modelle von automatischer Verlängerung möglich. In der Regel beträgt die Kündigungsfrist ein Jahr; weniger wird als nicht sinnvoll erachtet; langfristige Planung ist grundlegend. Vergleichbar ist die Situation bei FAST.

Im SLA können auch Konventionalstrafen vereinbart werden. Diese sind anerkanntermassen nicht in erster Linie als Entschädigung für allenfalls verlorene Daten zu interpretieren, sondern als Anreiz für den Anbieter, den Datenverlust zu vermeiden. FAST gibt aber zu bedenken, dass die Reputation des Anbieters Anreiz genug ist: Falls die Firma Daten z.B. eines Staatsarchivs verlieren würde, müsste sie zweifellos ihr Geschäft einstellen.

3.2 Zertifizierung: Wie stehen die Anbieter zur Diskussion über Trusted Digital Repositories und entsprechenden Zertifizierungsbemühungen?

COLT hat sich dem Ziel der Informationssicherheit verschrieben und orientiert sich dabei am ISO-Standard 17799⁶. Die Firma ist zertifiziert nach ISO 27001⁷, dem heute strengsten Standard für Informationssicherheits-Management. Dazu muss sie sämtliche Policies, Methoden, Technologien offen legen. Für bestimmte Kundengruppen gelten zudem spezielle Überprüfungen und Zertifizierungen. So untersucht die Eidgenössische Bankenkommission EBK die Angebote, welche COLT für Banken zur Verfügung stellt.

⁶ ISO/IEC 17799:2005, Information technology -- Security techniques -- Code of practice for information security management.

http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39612

⁷ ISO/IEC 27001:2005, Information technology -- Security techniques -- Information security management systems – Requirements.

http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=42103

Für FAST ist das Thema der Zertifizierung weniger wichtig, da die Firma kein Lösungsentwickler ist. Unter Umständen ist FAST auch gar kein Vertragspartner, sondern bietet seine Produkte über Drittfirmen an.

COLT weist auf die Zauberformel der Informationssicherheit hin: CIA, kurz für *confidentiality – integrity – availability* (Vertraulichkeit – Integrität – Verfügbarkeit). Diese drei Komponenten machen Informationssicherheit aus; sie werden aber von verschiedenen Kunden unterschiedlich gewichtet, was unterschiedliche Lösungen nötig macht. Die drei Komponenten sind teilweise gegenläufig; das bedeutet, dass die eine leichter erreicht werden kann, wenn eine andere weniger wichtig ist.

Die KOST stellt sich die Frage, ob sich die Archive eine Zertifizierung nach archivischen Kriterien⁸ überhaupt leisten können (oder leisten sollen, da die Anbieter bereits nach anderen, unter Umständen strengeren Kriterien zertifiziert sind). Die Anbieter relativieren: Die Zertifizierung sei bei weitem nicht der grösste Kostenblock; das Auditing müsse aber vom Leistungsbezügler, dem Archiv, erbracht werden.

3.3 Technische Fragen: Wie erfolgt der Zugriff auf die Daten?

FAST stellt zwei verschiedene APIs zur Verfügung (ein C-API und das S3-API von Amazon). *Content Addressed Storage* bedeutet, dass jedes eingelieferte Datenpaket einen SHA-Hashwert⁹ als Quittung erhält. Nach Authentifizierung und mit Vorweisen dieser Quittung kann auf das archivierte Datenpaket zugegriffen werden. Ein umfangreicher Export (zum Beispiel bei einem Anbieterwechsel) wird direkt im Rechenzentrum durchgeführt werden.

3.4 Kosten: Mit welcher Grössenordnung ist zu rechnen?

FAST hat ein Kostenbeispiel in der Grössenordnung von 0,15€ pro GB und Monat errechnet. COLT kann keine Kostenschätzung liefern, weil die Firma keine Systementwicklung macht, sondern für jeden Kunden die optimale Lösung erarbeitet. Es könnte allerdings eine Preisspanne angegeben werden (wird nachgeliefert).

3.5 Realisierbarkeit: Ist das überhaupt etwas für die Staatsarchive?

Die Schweizer Staatsarchive sind relativ kleine Kunden mit spezifischen Ansprüchen (Integrität ist absolut zentral, Verfügbarkeit sekundär, etc. – siehe dazu die Zusammenfassung der Einführungsveranstaltung). Die KOST vermutet, dass aus Kostengründen ausschliesslich eine Lösung ab Stange in Frage kommt. Zudem befürchtet sie, dass das auf die Finanzwelt zugeschnittene allgemeine Sicherheitslevel eines Anbieters wie COLT grundsätzlich zu teuer sein könnte.

Tatsächlich kann an der Grundinfrastruktur von COLT wenig geändert werden: Die Datacenter sind rund um die Uhr bedient, verfügen über eine unabhängige Stromversorgung, die physische Sicherheit wird von einer Drittfirma gewährleistet. Varianten sind möglich in der inneren Organisation (dedizierte Räume, gemietete Racks etc.). Infrastrukturkosten sind jedoch nicht der grösste Kostenpart.

⁸ RLG Report "Trusted Digital Repositories: Attributes and Responsibilities"
<http://www.oclc.org/programs/ourwork/past/trustedrep/repositories.pdf>

⁹ Secure Hash Algorithm <http://de.wikipedia.org/wiki/SHA-1>

Zu beachten ist, dass COLT keine fertigen Speicherplatzlösungen anbietet, sondern für jeden Kunden eine auf einer Analyse seiner Situation und Bedürfnisse beruhende Lösung erarbeitet. Es ist offensichtlich, dass dies für kleine Kunden den Preis in die Höhe treibt. Sowohl COLT als auch FAST raten den Archiven deshalb dringend, gegenüber dem Anbieter gemeinsam aufzutreten. Damit werden auch Lösungen mit dedizierter Infrastruktur machbar und sinnvoll.