KOST.07

Kolloquium „Archivtaugliche Speicherinfrastruktur“

Thema: Speichern im Netzwerk (verteilte Speicherung)

Inhalt

1 Einleitung

Als fünftes und letztes Speicherkonzept stellt die KOST in ihrem Kolloquium "Archivtaugliche Speicherinfrastruktur" unter dem Titel "Speichern im Netzwerk" die verteilte Speicherung vor. Dies ist ein Konzept, das in archivischen Kreisen noch kaum diskutiert, geschweige denn erprobt wird. Das Prinzip der *Peer-to-Peer*-Netzwerke (P2P) ist allerdings in der Informatik verbreitet und bewährt. Als Input wurden zwei Referenten aus dem ausserarchivischen Bereich eingeladen. Simon Margulies stellte das Projekt distarnet[1] des imaging & media lab (IML) der Universität Basel[2] vor, welches ein Protokoll zur verteilten Speicherung erarbeitet hat. Jan Dirk Brinksma von der Bibliothek für Exakte Wissenschaften (BEWI) der Universität Bern[3] berichtete über seine Erfahrungen mit LOCKSS[4], einem P2P-Netzwerk für e-Journals. Aus der Sicht der KOST war anhand dieser Referate und der darauffolgenden Diskussion zu klären, ob der Ansatz der verteilten Speicherung für die Staatsarchive eine Alternative zu anderen, besser bekannten Ansätzen darstellen könnte.

2 Referate

2.1 Simon Margulies, Projekt distarnet, imaging & media lab der Universität Basel

Distarnet ist ein vom Schweizerischen Nationalfonds unterstütztes Forschungsprojekt am IML, das eine Lösung für die Probleme Datenträgermigration und Bitstrom-Erhalt erarbeitet. Dazu wurde ein Protokoll für ein verteiltes System definiert, das über ein Netzwerk kommuniziert, dem Benutzer gegenüber aber transparent erscheint. Darin ist Distarnet anderen in diesem Kolloquium untersuchten Ansätzen, z.B. der Blackbox, nicht unähnlich. Durch weitgehende Automatisierung sollen die Integrität, Datensicherheit und Durchführbarkeit gewährleistet werden. Als hauptsächliche Risiken, vor denen die Lösung schützen kann, sind der Ausfall eines Partners sowie der Verlust einzelner Dateien.

Distarnet ist ein *Peer-to-Peer*-Netzwerk, d.h. ein über ein Netzwerk verbundener, Zusammenschluss verschiedener Partner ("Knoten"), die alle gleichberechtigt sind. Durch den Verzicht auf eine zentrale Instanz ist weitgehende Redundanz und Ausfallsicherheit gewährleistet. Digitale Objekte werden jeweils von einem Knoten unter Angabe der nötigen Redundanz ins Netzwerk eingespiesen. Ein Evaluationsprozess kontrolliert die Redundanz. Dabei werden die einzelnen Knoten nach einer Kriterienliste bewertet; zur Bewertung tragen bei die geografische Distanz, der vorhandene Speicherplatz, die Geschwindigkeit der Verbindung, die Verfügbarkeit sowie (als wichtigstes Kriterium) die Frage, ob es sich um den ursprünglich einspeisenden Knoten handelt. Nach dieser Bewertung werden eine oder mehrere Kopien auf die bestplatzierten Knoten verteilt. Die Metainformationen über die archivierten Objekte befinden sich in einem ebenfalls verteilten Informationsspeicher,

[1] http://www.distarnet.ch/
[2] http://www.abmt.unibas.ch/
[3] http://www.bewi.unibe.ch/
[4] http://www.lockss.org/

der bei allen Knoten vorhanden ist. Dazu wird Kademlia[5] verwendet, eine *Distributed Hash Table* (DHT, verteilte Hash-Tabelle), die auch in vielen Filesharing-Tools zur Anwendung kommt.

Die Evaluation ist ein kontinuierlicher Prozess: Der gesamte Archivbestand wird regelmässig evaluiert; dabei werden zu kleine oder zu grosse Redundanzen ausgeglichen. Damit ist einerseits die Datensicherheit im Distarnet gewährleistet. Anderseits deckt dieser Prozess auch die Datenträgermigration ab: An einem Knoten kann das Speichermedium im laufenden Betrieb ersetzt werden; die nötige Redundanz wird danach automatisch wieder hergestellt, indem die zwischenzeitlich "verlorenen" Daten wieder kopiert werden.

Die Integrität der archivierten Objekte wird in periodischen Intervallen über Hash-Werte kontrolliert. So findet auch vor jedem Kopiervorgang eine Hash-Überprüfung statt. Da die Hash-Werte in der DHT und somit bei allen Knoten hinterlegt sind, können Konflikte durch unterschiedliche Versionen eines Objekts an verschiedenen Knoten einfach gelöst werden, indem die korrupte Version identifiziert wird.

Damit ein Distarnet-Speichernetzwerk funktioniert, muss das verwendete Netzwerk hinreichend stabil sein, da die zeitweise Nicht-Erreichbarkeit einzelner Knoten die automatische Duplizierung der dort gespeicherten Objekte auslöst. Zudem sollten die einzelnen Knoten eine ungefähr gleich grosse Datenmenge einspeisen.

Jeder Knoten besitzt im Normalfall eine Kopie der von ihm eingespiesenen Daten. Zu diesen Daten gewährt der Knoten Zugang nach aussen. Die Kopien der Daten anderer Knoten, die ebenfalls dort lagern, sind nicht sichtbar.

Im Rahmen der Projektarbeit wurde das Protokoll entwickelt und lokal getestet. Allerdings konnte keine geografisch verteilte Implementierung mit mehreren Teilnehmern getestet werden. Deshalb wurden auch keine Kostenmodelle entwickelt; anzumerken ist allerdings, dass Distarnet ursprünglich den Anspruch erhob, durch Automatisierung und billige Hardware besonders kostengünstig zu sein. Die Zukunft von Distarnet ist zurzeit noch nicht klar; das Projekt wird eben abgeschlossen.

2.2 Jan Dirk Brinksma, Bibliothek der Exakten Wissenschaften, Universität Bern

In der Bibliothekswelt ist man mit LOCKSS, einem System zur Archivierung von e-Journals, bereits mehrere Schritte weiter. Das System wurde ursprünglich an der Stanford University entwickelt und ist inzwischen in 200 akademischen Bibliotheken (in Zusammenarbeit mit 200 Verlagen) weltweit im Einsatz. Das Schwergewicht liegt in Amerika; in der Schweiz hat Jan Dirk Brinksma von der Bibliothek der Exakten Wissenschaften der Uni Bern als erster den Einsatz von LOCKSS erprobt. Er berichtete von seinen Erfahrungen als User.

LOCKSS steht für *"Lots Of Copies Keep Stuff Safe"* und weist darauf hin, dass damit ein aus der analogen Bibliothekswelt bekanntes Prinzip auf die digitale Welt und ihre spezifische Problemlage übertragen wurde: Indem Kopien elektronischer Zeitschriften bei allen an LOCKSS beteiligten Bibliotheken aufbewahrt werden, wird die

[5] http://de.wikipedia.org/wiki/Kademlia; siehe dazu auch Maymounkov, Petar; Mazières, David; Kademlia: A Peer-to-peer Information System Based on the XOR Metric. http://pdos.csail.mit.edu/~petar/papers/maymounkov-kademlia-lncs.pdf

Abhängigkeit von der einen ursprünglichen Kopie des Verlags gemindert. LOCKSS besteht aus einem Netzwerk von Knoten in Bibliotheken, sogenannten LOCKSS-Boxen, die mit Anbietern von Open-Access-Zeitschriften und traditionellen akademischen Verlagen verbunden sind. Die LOCKSS-Box ruft die Zeitschriften wie ein Web-Crawler bei den entsprechend freigeschalteten Verlagen ab und speichert sie lokal. Beim Zugang funktioniert sie als Web-Proxy: Wenn ein Artikel angefordert wird, leitet die Box die Anfrage an den Verlag weiter und gibt den Artikel dem Benutzer aus. Nur wenn der Artikel am originalen Ort nicht mehr verfügbar ist, greift die Box auf die lokale Kopie zurück. Ein robuster Abstimmungsmechanismus sorgt in regelmässigen Abständen dafür, dass auf allen dem Netz angeschlossenen LOCKSS-Boxen authentische Kopien der digitalen Objekte gespeichert sind.

LOCKSS wurde mit besonderer Beachtung der Anwenderfreundlichkeit entwickelt. Herr Brinksma berichtet, dass sich die Installation in der Tat als völlig problemlos herausgestellt hat. Benötigt wird ein normaler Computer mit Netzanschluss, auf welchem ein schlankes Open-Source-Betriebssystem (OpenBSD[6]) sowie die frei erhältliche LOCKSS-Software laufen. Allerdings muss die Implementation in der BEWI als Testinstallation gelten, da sie sich auf das Speichern von Open-Access-Zeitschriften aus den Geistes- und Sozialwissenschaften beschränkt und nicht als Proxy beim Lesen von Zeitschriften zwischengeschaltet ist.

Neben diesem Standardform von LOCKSS sind zwei Sonderformen von Interesse: CLOCKSS (*Controlled LOCKSS*)[7], eine Zusammenarbeit von sieben Bibliotheken und 11 Verlagen, baut ein Dark Archive von wissenschaftlichen Publikationen auf, um diese für den Fall für die Öffentlichkeit zu erhalten, dass ein Verlag seine Tätigkeit einstellt. Private LOCKSS ist ein LOCKSS-Netzwerk für eine kontrollierte Benutzergruppe mit beschränktem Zugang. Dieser Ansatz könnte für Archive interessant sein.

3 Diskussion

3.1 Wie könnten Staatsarchive ein verteiltes Speichernetzwerk realisieren?

Ein P2P-Netzwerk zur Datenspeicherung ist für die Staatsarchive grundsätzlich in zwei Varianten denkbar:

In der ersten Variante schliessen sich mehrere Archive zu einem Netzwerk zusammen. In der LOCKSS-Terminologie ist jedes Archiv damit gleichzeitig Verlag und Knoten: Es speist seine eigenen Unterlagen in das Netzwerk ein; Kopien werden dann an verschiedenen anderen Knoten gespeichert. Für diese Lösung spricht einiges: Kooperationen zwischen einzelnen Staatsarchiven existieren bereits, zudem ist so das wichtige Postulat der geografischen Redundanz für die Speicherung ideal erfüllt. Als problematisch wird hingegen der Datenschutz angesehen: Können Archive ihre Unterlagen (die Schutzfristen unterstehen und möglicherweise besonders sensibel sind) an einem Ort speichern, der nicht unter ihrer Kontrolle ist? Müssen sie diese

[6] http://de.wikipedia.org/wiki/OpenBSD; siehe auch http://www.openbsd.org/.
[7] http://www.clockss.org/

verschlüsseln und die damit verbundenen Risiken in Kauf nehmen[8]? Oder wäre es denkbar, dass über Verträge das Vertrauen in die geschützte Aufbewahrung bei Partnerarchiven gebildet werden kann? Es herrscht jedenfalls Einigkeit, dass dieser Ansatz zwar vielversprechend, aber mit Problemen organisatorischer und psychologischer Art verbunden ist.

Die zweite für Archive mögliche Variante ist ein internes Netzwerk, d.h. die Vernetzung mehrerer Speichergefässe mittels des Distarnet-Protokolls. Hier handelt es sich um eine weitergehende Anwendung der vorgestellten Ideen. Die archivierten Daten müssen auf jeden Fall redundant gespeichert werden. Das Distarnet-Protokoll ist ein möglicher Mechanismus, diese Redundanz und damit die Datenpersistenz zu kontrollieren. Dabei darf nicht vergessen werden, dass P2P-Netzwerke bereits bestens bekannt und bewährt sind. So beruhen sehr populäre Programme zum Dateitausch über das Internet auf der P2P-Technologie (Napster, Gnutella, BitTorrent, I2P etc.)[9]. Die dort verwendeten Protokolle und Mechanismen können als stabil und bestens getestet gelten. Unter anderem wird z.B. die Koexistenz verschiedener Software-Versionen gewährleistet. Wenn sich Archive für die Anwendung der P2P-Technologie für die Datenspeicherung interessieren, sollten sie unbedingt das Potential dieser Netzwerke untersuchen und nutzbar machen, indem sie es auf ihre Bedürfnisse adaptieren.

3.2 Wie werden archivische Anforderungen erfüllt?

Es erweist sich als schwierig, die Erfüllung archivischer Anforderungen zu untersuchen, da wir sie an keinem realen System testen können. (LOCKSS ist momentan sehr stark auf die Archivierung von e-Journals zugeschnitten, so dass dort gewonnene Erkenntnisse kaum auf die Anforderungen von Staatsarchiven angepasst werden können.) Grundsätzlich können aber ein paar Punkte festgehalten werden:

– Durch die Absenz eines Single Point of Failure, also eines zentralen Zugriffs- und Verwaltungsortes, vergrössert die verteilte Speicherung die Datensicherheit.

– Der Aufbau eines archivinternen Speichernetzwerks, das vom Distarnet-Protokoll oder einem anderen, offengelegten Protokoll geregelt wird, ermöglicht den Archiven eine deutlich grössere intellektuelle Kontrolle über die Speichersoftware als sie etwa bei einer Blackbox-Lösung möglich ist. Zudem sind Speicherlogik und Hardware sauber getrennt

3.3 Schlussfolgerungen

Eine wichtige Schlussfolgerung der Veranstaltung ist, dass das Speichern im Netzwerk für die Staatsarchive im Moment noch nicht einsatzfähig ist. Es existieren bisher keine (Test-)Implementationen, welche die Machbarkeit realistisch darstellen und Kostenschätzungen ermöglichen. Eine P2P-Lösung hätte, wenn die organisatorischen und gefühlsmässigen Probleme geklärt sind, ein grosses Potential für die Zusammenarbeit einzelner Archive.

---

[8] Es steht ausser Frage, dass die Kommunikation zwischen den Knoten über das öffentliche Netzwerk verschlüsselt sein muss. Problematisch ist einzig die verschlüsselte Aufbewahrung.

[9] http://de.wikipedia.org/wiki/File_Sharing