



BABS

Bibliothekarische Archivierungs- und Bereitstellungssystem

Dr. Bernd Reiner (reiner@lrz.de)

Veröffentlichungen in digitaler Form nehmen im Wissenschaftsbetrieb wie auch im gesellschaftlichen Leben insgesamt einen immer höheren Stellenwert ein. Oftmals wird, wie z.B. bei Dissertationen und bei amtlichen Publikationen, auf ein gedrucktes Pendant ganz verzichtet. Während die Digitalisierung dem Nutzer den Zugang und den Umgang mit der Information beschleunigt und insgesamt erleichtert, entstehen den Bibliotheken weltweit dadurch, sowohl aus organisatorischer und rechtlicher, als auch aus technischer Sicht, neue Herausforderungen. Zusätzlich sind die Bibliotheken mit einer jährlich stark steigenden Zahl von digitalen Objekten konfrontiert. Die Objekte sollen nicht nur verwaltet und gespeichert, sondern auch langfristig zugänglich gemacht werden. Diese Aufgabe wird erschwert durch den raschen technologischen Wandel im Bereich der Hard- und Software und durch die natürlichen physikalischen Grenzen der Datenträger.

Aus diesem Grund wurde die bereits 2004 begonnene Kooperation zwischen der Bayerischen Staatsbibliothek (BSB) und dem Leibniz-Rechenzentrum im Juli 2005 durch ein Projekt im Bereich der Langzeitarchivierung (LZA) ausgeweitet. Dieses Projekt wird durch die Deutsche Forschungsgemeinschaft (DFG) gefördert und trägt den Namen BABS (Bibliothekarische Archivierungs- und Bereitstellungssystem; www.babs-muenchen.de). Das Projekt BABS ist zunächst auf eine Dauer von zwei Jahren ausgelegt. Das Ziel ist der exemplarische Aufbau einer organisatorischen und technischen Infrastruktur für die Langzeitarchivierung von Netzpublikationen und von sonstigen E-Medien unterschiedlicher Provenienz. Die gewonnene Erfahrung soll zur Etablierung neuer Workflows und Verfahren des Datenmanagements in der BSB und dem LRZ beitragen und das Fundament einer langfristig tragfähigen Archivierungs- und Bereitstellungsumgebung werden.

Innerhalb des Projektes wird sich die BSB verstärkt mit der Sammlung, Erschließung (Ingest), Verwaltung (Data Management) und Zugriffsteuerung (Access) beschäftigen, während sich das LRZ mit der Archivierung im eigentlichen Sinne (Archival Storage) und den Erhaltungsmaßnahmen (Preservation Planning) befasst. Die Administration der entsprechenden Komponenten wird von dem jeweiligen Projektpartner übernommen. Damit werden alle wichtigen Bereiche eines langfristig funktionierenden Archives abgedeckt und folgen somit dem allgemeinen Open Archival Information Systems (OAIS) Referenz Modell. Abbildung 1 zeigt das Archivierungs- und Bereitstellungssystem BABS mit den OAIS Prozessen Ingest, Data Management, Archival Storage, Access und Preservation Planning.

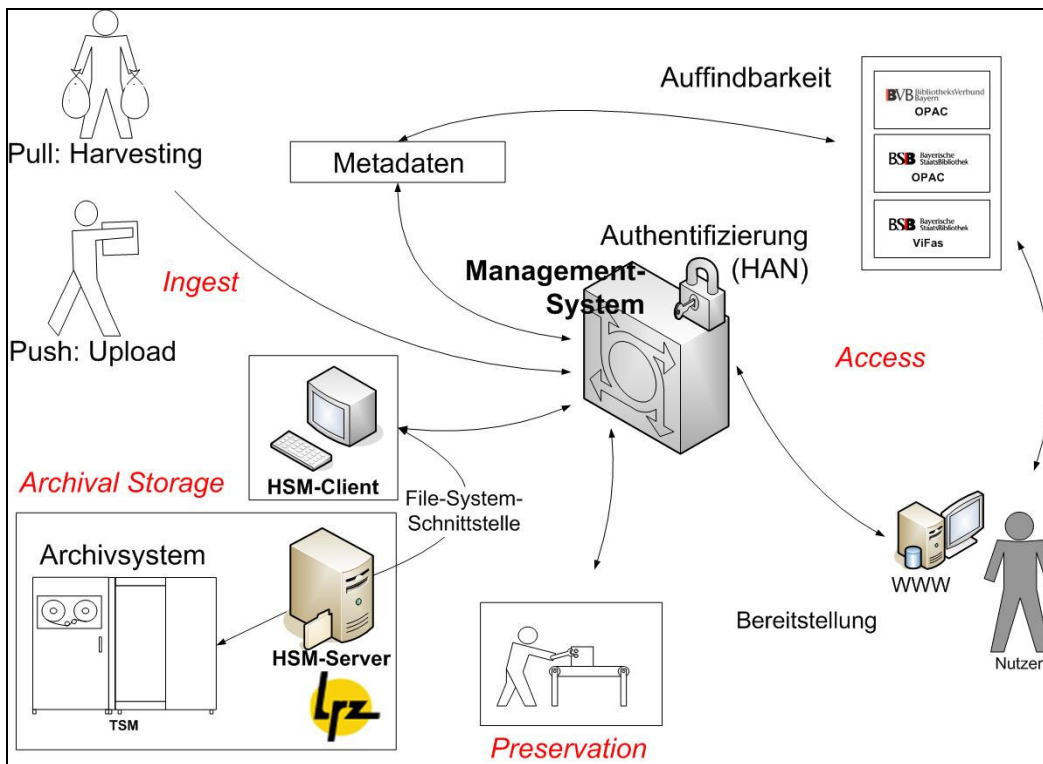


Abbildung 1: Archivierungs- und Bereitstellungssystem BABS

Abbildung 1 zeigt schematisch das im Pilotprojekt entwickelte funktionale Konzept eines Archivierungs- und Bereitstellungssystems für heterogene Netzpublikationen verschiedener Provenienz:

- Die Objekte werden eingesammelt oder manuell eingestellt (→ „Harvest“ bzw. „Upload“) und in das Management-System eingespeist (→ „Ingest“).
- Die Verwaltung der Objekte wird im Management-System erfolgen. Die bibliographischen Metadaten werden aus dem lokalen Katalog bzw. den Internetquellendatenbanken übernommen; technische Metadaten aus den Dateien extrahiert; administrative, die Prozesse beschreibende Metadaten werden automatisch protokolliert.
- Über eine leistungsfähige Schnittstelle (→ „File-System-Schnittstelle“) werden die Daten zu dem Archiv- und Backup-System am LRZ (→ „Archivsystem“) übertragen und dort archiviert.
- Die Recherche des Nutzers findet über WWW, entweder über das Lokalsystem bzw. die ViFas oder über eine eigene Suchmaske statt (→ „Access“) und löst für die Bereitstellung einen Rückholungsprozess aus dem Archivspeicher aus.
- Aufgrund der Rechtslage oder bestimmter Nutzungsvereinbarungen erfordern manche Archivobjekte einen kontrollierten Zugriff mit abgestuften Zugriffsrechten (z. B. nur in der Bibliothek, nur einem eingeschränkten Nutzerkreis, erst nach Ablauf von Sperrfristen) (→ „Authentifizierung“).
- Notwendige Erhaltungsmaßnahmen (z.B. Migrationen von Dateien, deren Formate von Veralterung bedroht sind) werden bei Bedarf durchgeführt (→ „Preservation“).

Zur Speicherung der Daten (Archival Storage) wird das am LRZ betriebene Archiv- und Backupsystem unter dem Softwarepaket Tivoli Storage Manager (TSM) der Firma IBM verwendet.

TSM ist als Client-Serverapplikation für heterogene Rechnerlandschaften konzipiert und basiert auf offenen Standards, wie dem IEEE Mass Storage Reference Modell. Der TSM Serveranteil wird im Projekt die Funktion eines Archivspeichersystems, speziell für die Langzeitarchivierung, übernehmen. Das System verfügt intern über essentielle Funktionalitäten, die für Langzeitarchivsysteme als Voraussetzung gelten:

- Automatische und zentrale Verwaltung aller gespeicherten Daten
- Einfache Administrierbarkeit
- Zugriffskontrolle und Zugriffsbeschränkung
- Fehlerüberwachung der Datenträger
- Recycling der Datenträger bei auftretender Fragmentierung
- Migrationsmechanismen
- Automatisierte Sicherung (Backup) von Daten
- Redundante Datenhaltung (z.B. zusätzliches Backup von Archivdaten)
- Keine Einschränkung der verwendbaren Dateiformate (Backup & Archivierung)
- Verwaltung der Daten und Metadaten durch ein Datenbankmanagementsystem
- Konsistente Datenhaltung
- Transaktionsverwaltung und Recoverymöglichkeiten
- Sicherungskonzepte für das Datenbankmanagementsystem (RAID 5 + Spiegelung auf Festplatten + Backup auf Magnetbänder)
- Validierte Datenübertragung vom Client zum Server (CRC, cyclic redundancy check)
- Kompression der Daten (Übertragung und/oder Speicherung)
- Gute Skalierbarkeit des Gesamtsystems

Bei der Realisierung des Bereitstellungssystems BABS können zur Anbindung des Archivspeichersystems (Magnetbandroboter) an die LZA-Verwaltungskomponente im Wesentlichen drei unterschiedliche TSM Schnittstellen verwendet werden:

- Standard Client-Schnittstelle
- Application Programming Interface (API)
- Hierarchisches Dateisystem (Hierarchisches Speicher Management System)

Die Standard-Schnittstelle erlaubt die Archivierung (und Backup) von Objekten unter Angabe zusätzlicher Informationen (Metadaten), ist allerdings relativ unflexibel. Das Application Programming Interface ist aufgrund der freien Programmierbarkeit sehr flexibel, allerdings mit einem sehr hohen Arbeitsaufwand verbunden und hat einen weiteren Nachteil: Dateien, die über die API archiviert wurden, können nur über die API und nicht zusätzlich über die Standard Client-Schnittstelle angefragt werden. Dies widerspricht der LZA-Forderung nach generischer Rekonstruktion, da die Dateien nur über das implementierte Managementsystem zugegriffen werden können. Bei der Anbindung über ein hierarchisches Dateisystem werden die weithin etablierten Dateisystem-Mechanismen für eine transparente Anbindung an Hierarchien von Speichermedien genutzt. Die gesamte Komplexität der Speicherhierarchie bleibt vor der Applikation verborgen. Ein HSM-System (Hierarchisches Speicher Management System) erweitert den Festplattenspeicher, um die Kapazität einer Vielzahl tertiärer Speichermedien und kann als normales Dateisystem mit nahezu unbegrenzter Speicherkapazität angesehen werden. In Wirklichkeit ist das virtuelle Dateisystem eines HSM-Systems, unterteilt in einen begrenzten Festplattencache auf dem der Anwender, bzw. die Anwendung arbeitet (Schreiben und Lesen von Daten) und einem angegliederten Tertiärspeichersystem, mit einer robotergesteuerten Medienbibliothek. Außerdem kann durch die Verwendung der hierarchischen Dateisystemschnittstelle auf proprietäre TSM Schnittstellen verzichtet werden. Aufgrund der hier aufgezeigten Vorteile fiel die Wahl für dieses Projekt auf die hierarchische Dateisystemschnittstelle des TSM Systems.

Abbildung 2 zeigt die Systemarchitektur des in diesem Projekt zu entwickelnden Bereitstellungs- und Archivierungssystems BABS. Die Anbindung des LZA-Managementsystems der BSB an das HSM-System des LRZ erfolgt über das Münchener Wissenschaftsnetz (MWN) mittels NFS (Network File Service) oder scopy (secure copy). Im unteren Bereich der Abbildung 2 lässt sich die Speicherhierarchie erkennen, bestehend aus einem Festplattensystem des HSM-Clients, einem Festplattensystem (RAID) des HSM-Servers und dem robotergesteuerten Tertiärspeichersystem. Die Trennung des HSM-Systems in Client und Server folgt der von Tivoli Storage Manager realisierten Architektur. Ein HSM-Server kann somit mehrere HSM-Clients bedienen. Ein HSM-System erweitert die bereits bestehende Speicherhierarchie der Primär- und Sekundärspeicher um die Komponente der Tertiärspeicher.

Da das Festplattensystem des HSM-Clients über das Münchener Wissenschaftsnetz (MWN) mittels NFS an das LZA-Managementsystem der BSB angebunden ist, kann dieser Speicherplatz als lokale Festplatte behandelt werden. Beim Einfügen von Daten werden die Dateien auf der „lokalen“ NFS-Festplatte des HSM-Clients abgelegt (siehe Abbildung 2). Der HSM-Client verschiebt die Dateien entsprechend den eingestellten Parametern zum HSM-Server. Dieser wiederum migriert die Dateien auf Magnetbänder. Zur Steigerung der Zugriffperformance werden die Festplattensysteme des HSM-Clients und des HSM-Servers als Cache genutzt. Dies bedeutet, dass Dateien, die auf Magnetband migriert wurden erst von den Plattenspeichern verdrängt werden, wenn neuer Speicherplatz benötigt wird. Diese Dateien werden auch als vormigrierte Dateien bezeichnet. Erst bei einem direkten Zugriff (z.B. Lesezugriff) auf eine Datei wird diese, falls sie sich nicht im Cache befindet, vom Magnetband geladen. Nur ein etwaiger längerer Ladevorgang lässt auf einen Tertiärspeicherzugriff schließen. Die gesamte Komplexität des Auslagerungs- bzw. des Ladevorganges der Dateien bleibt vor dem LZA-Managementsystem verborgen.

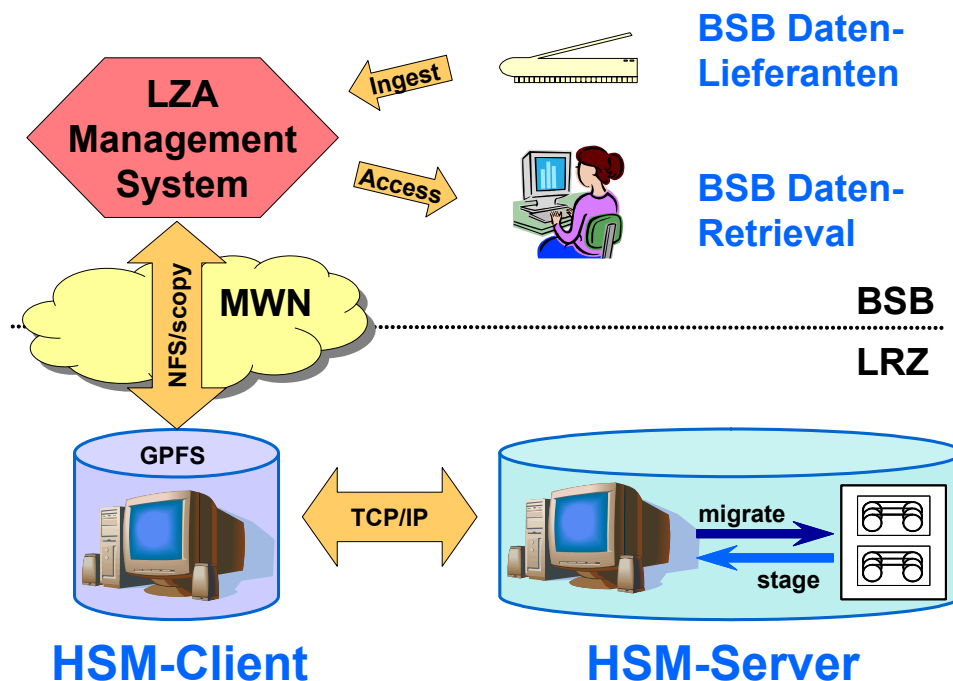


Abbildung 2: Systemarchitektur des Bereitstellungs- und Archivsystems BABS

Sowohl der HSM-Client (CPU: 2 x 3 GHz Intel Xeon; Main-Memory: 4 GByte) als auch der HSM-Server (CPU: 2 x 3 GHz Intel Xeon; Main-Memory: 4 GByte) wurden unter Suse Linux Enterprise Server (SLES) in der Version 8 installiert. Für den HSM-Server wurde die TSM Version 5.2 verwendet. Aufgrund der IBM Supportmatrix konnte hier noch nicht auf die neuste TSM Version 5.3 zurückgegriffen werden. Bei der Installation des HSM-Clients wurde die neueste TSM Version 5.3 eingesetzt. Um die HSM Funktionalität zu nutzen musste zusätzlich auf dem HSM-Client das General Parallel File System (GPFS) von IBM in der Version 2.3.0.5 aufgesetzt werden. Anfängliche Probleme beim NFS-Zugang konnten durch die im November 2005 erschienene TSM Version 5.3.2.1 und dem neusten Service Level des GPFS in der Version 2.3.0.7 beseitigt werden.

Durch die Wahl geeigneter Parameter bei der Konfiguration von NFS, GPFS (General Parallel File System) und des HSM-Systems konnte die Performance verbessert und die Funktionsweise auf die Anforderungen angepasst werden. Tests der Schnittstelle bei mittleren Datenmengen, im Bereich von bis zu 600 GByte, wurden erfolgreich durchgeführt. Sowohl der Datendurchsatz (Sequentielles Lesen, bzw. sequentielles Schreiben: 45 - 54 MByte/s), als auch die Funktionsweise des HSM-Systems versprechen eine gute Skalierbarkeit des Archivierungssystems. Wie sich die Skalierbarkeit des Gesamtsystems darstellt werden weitere Tests zeigen. Insgesamt stellt das hier vorgestellte Bereitstellungs- und Archivierungssystem BABS eine flexible Lösung für die Langzeitarchivierung dar.