

Bundesarchiv

Das Bundesarchiv digitalisiert immer mehr seiner Dokumente und steht dabei vor großen Aufgaben. Jährlich fällt ein hohes Volumen an digitalisierten Dateien an, das evaluiert und in das Archiv übernommen werden muss. Dabei helfen auch externe Digitalisierungsdienstleister. Auf der Suche nach Wegen, Dateien schneller ins Archiv übernehmen zu können, stieß

die Bundesbehörde auf die Silent Bricks von FAST LTA. Nach einem besonders schnellen Einführungsprojekt sind mittlerweile 70 der hochperformanten Systeme aus SSD-Speichern und transportierbaren Storage-Containern im Einsatz, um digitalisierte Daten innerhalb kürzester Zeit ins Koblenzer Archiv zu bringen.



„Mit den Silent Bricks kommen wir auf Datentransferraten von 500 bis 700 Megabyte pro Sekunde bei großen Dateien, statt vorher 80 MB: Damit hat sich ein immenser Vorteil und eine erhebliche Prozessverschlankeung ergeben. Wir konnten die Technologie ohne jede Anpassung sehr einfach implementieren und an externe Dienstleister ausrollen“

Timo Dommermuth, IT-Spezialist beim Bundesarchiv

Zehnfach höhere Geschwindigkeit mit Silent Bricks:

Wie das Bundesarchiv das Gedächtnis der Gesellschaft digitalisiert

Die Wochenschauen aus der Kriegszeit, Filme wie „Rosen für den Staatsanwalt“, Dokumente seit Ende des 14. Jahrhunderts aufwärts und seit 2021 auch die Stasi-Unterlagen: Im Bundesarchiv lagert das Gedächtnis unserer Gesellschaft. Die Dokumente auf Papier oder Film werden immer weiter digitalisiert: Ab 2024 sollen es bis zu 80 Millionen Seiten im Jahr sein. Dabei spielen die Silent Bricks von FAST LTA als schnelles und transportables Speichermedium eine zentrale Rolle. Die Geschwindigkeit der Digitalisierung konnte damit

im Vergleich zur vorherigen Lösung fast verzehnfacht werden. Durch den Technologiewechsel stehen die Daten doppelt so schnell im Archiv zur Verfügung. Das Bundesarchiv und seine Digitalisierungsdienstleister speichern die Dateien zunächst auf den mobilen, praktisch unverwüchtlichen Speichermedien, die nach Koblenz transportiert und dort auf Magnetband für die Ewigkeit aufbewahrt werden. Bei jeder neuen Aktenanfrage gilt mittlerweile das Prinzip „Digitalisierung on Demand“ – jede dafür angefasste Papierakte wird digitalisiert.

HERAUSFORDERUNGEN

1

Bis zu zehn Petabyte sollen pro Jahr digitalisiert werden

2

Traditionelle Festplatten erwiesen sich als Flaschenhals

3

Sicherer Transport großer Datenmengen über längere Distanzen hinweg

Jede Zeit bringt neue Erkenntnisse: Umso wichtiger ist es, Dokumente der Vergangenheit immer wieder aus dem aktuellen Zeitgeist und Kenntnisstand heraus bewerten zu können. Archive wie das Bundesarchiv sind dafür essentiell. „Jedes Archiv ist ein Ort der Vergewisserung, die individuelle und kollektive Erinnerung möglich macht. Im Zeitalter der Information und des digitalen Wandels sind Archive Garanten der Volkssouveränität“, so formuliert es Prof. Dr. Michael Hollmann, Präsident des Bundesarchivs.

Einfach gigantisch viele Daten

Das Bundesarchiv hat den gesetzlichen Auftrag, das Archivgut des Bundes auf Dauer zu sichern und nutzbar zu machen. Aufbewahrt werden Akten, Karten, Bilder, Plakate, Filme und Tonaufzeichnungen in analoger und digitaler Form. Um die 2300 Mitarbeitenden sind derzeit im Bundesarchiv tätig, etwa 100 davon in der IT. Die Digitalisierung ist eine echte Herkulesaufgabe, jedes Jahr kommen zwischen fünf und zehn Petabyte an Speichervolumen hinzu. Es gibt viele analoge Quellen, Ende 2020 waren es beispielsweise allein

428 laufende Kilometer Schriftgut und über eine Million Filmrollen. Seit Mitte 2021 sind auch die Stasi-Unterlagen der Gauck-Behörde in die Verantwortung des Bundesarchivs übergegangen: Hier entsteht weiterer Digitalisierungsbedarf. Zudem werden alle Filme, die in Deutschland von der Filmförderung profitieren, im Bundesarchiv gespeichert. Die meisten Inhalte sind allerdings Akten, darunter zum Beispiel Protokolle des Bundeskabinetts und von Behörden. Die archivierten Dokumente reichen bis zurück ins Heilige Römische Reich Deutscher Nation. Auch geheime Unterlagen lagern im Bundesarchiv.

Was tun gegen Bottlenecks im Digitalisierungsprozess?

Die ersten Überlegungen zum Projekt machte man sich im Herbst 2018. „Wir standen vor dem Problem, dass wir sehr viele Daten von externen Dienstleistern bekommen, Tendenz steigend. Damals kamen diese Daten auf USB-Medien oder -Festplatten zu uns und die Übernahme ging mit sehr viel Aufwand einher“, berichtet Timo Dommermuth, Referatsleiter IT-Betrieb beim Bundesarchiv. Zunächst mussten die Daten auf einen speziellen Speicherbereich herunter kopiert werden, um dann für die nächsten Weiterverarbeitungsschritte verfügbar zu sein. „Das Kopieren von diesen Datenträgern, meist USB-Festplatten, funktionierte nur mit 80 Megabyte pro Sekunde – im besten Fall. Das hat die Verarbeitung erheblich in die Länge gezogen, denn wir verarbeiten ja nicht nur einige Terabyte, sondern bis



zu zehn Petabyte im Jahr“, erklärt Dommermuth, der für das Projekt verantwortlich ist. Immerhin rund zwei Wochen dauerte es, bis die Daten auf so einer Festplatte entsprechend validiert und vollständig im Archiv ingestiert, also aufgenommen, waren. Denn aus dem dezidierten Speicherbereich heraus erfolgte eine Validierung und Transformation in das TAR-Format. Diese Datenpakete gelangten anschließend auf den Band-Archivspeicher. Die Überlegung bestand also darin, wie sich dieser Prozess wesentlich verschlanken, vereinfachen und beschleunigen lassen könnte.

Kopiervorgang übersprungen und Transferraten deutlich beschleunigt

Das war nicht ganz einfach, da es praktisch keine sinnvoll erscheinende Lösung am Markt zu geben schien. Schließlich erinnerten sich die IT-Experten an eine Präsentation der FAST LTA-Speicher, die sie vor einiger Zeit gesehen hatten. Damals war es um das klassische Konzept der Silent Bricks als Archiv-Lösung gegangen, die als energieeffizienter Cold Storage fungieren können. Bei den Datenmengen, mit denen es das Bundesarchiv zu tun hat, wäre das jedoch utopisch gewesen, meint Timo Dommermuth. Doch jetzt hatten die Experten eine zündende Idee: Wie wäre es, die Technologie einfach für den Austausch mit den Dienstleistern zu nutzen? „Wir haben uns an die Information erinnert, dass sich die Silent Bricks direkt in Controller schieben und mounten lassen. So könnten der Ingest-Vorgang und das Kopieren deutlich schneller werden“, fasst Dommermuth zusammen. Ende 2018 startete man einen Proof of Concept (POC) gemeinsam mit IT-Dienstleister und FAST LTA-Partner Bechtle AG. Gestartet wurde mit Silent Bricks HD in zwei Größen und einer SSD-Variante sowie den dazugehörigen Controllern. „Im POC haben wir festgestellt, dass wir mit den Silent Bricks Transferraten von 500 bis 700 Megabyte pro Sekunde, bei

großen Dateiformaten, hinbekommen: Damit hat sich ein immenser Vorteil und eine erhebliche Prozessverschlangung ergeben“, sagt Dommermuth. Die Durchlaufzeiten konnten mehr als halbiert werden, entsprechend schneller stehen die Informationen nun digital im Archiv zur Verfügung. Die Daten werden von den Dienstleistern direkt auf die Silent Bricks geschrieben, nach dem Transport im Bundesarchiv lediglich in die Controller geschoben und per NFS eingebunden: Der Kopiervorgang fällt komplett weg. Im ersten Projektschritt wurden dann zwanzig Silent Bricks angeschafft und zwei Controller, die jeweils fünf Einschübe für fünf Silent Bricks bieten. Wegen der höheren Performance fiel die Wahl auf die SSD-Variante.

Digitalisierungspartner einfach eingebunden

Schon im POC wurde mit einigen Dienstleistern, denen man gemeinsam mit Bechtle die Aufgabenstellung an die Hand gegeben hatte, die neue Speichertechnologie erprobt. „Wir konnten die FAST LTA-Technologie ohne Anpassungen sozusagen von der Stange verwenden und unsere Digitalisierungspartner haben uns zurückgemeldet, dass sie direkt reibungslos damit arbeiten konnten“, erinnert sich der IT-Spezialist. Vom ersten Gedanken bis zur Umsetzung habe es nur ein halbes Jahr gedauert. „Das ist für ein IT-Projekt mit einer Technologie, die zuvor noch nicht im Einsatz war, sehr zügig“, so Dommermuth. Etwa ein Drittel der Digitalisierung erledigt das Team des Bundesarchivs selbst, den Rest externe Dienstleister. Am Standort Berlin-Lichterfelde werden selbst Silent Bricks mit neuen Daten beschrieben und alle zwei Wochen per Regeltransport mit Koblenz ausgetauscht. Dafür kommen die besonders sicheren Transportbehälter zum Einsatz. „Wir haben auf das richtige Pferd gesetzt“, bilanziert Timo Dommermuth, „der Kontakt mit Bechtle ist tadellos und der Support von FAST LTA läuft sehr gut, beide Partner sind sehr flexibel. Probleme sind bisher nicht aufgetaucht“.

UMSETZUNG

1

Erfolgreicher Proof of Concept mit Einbindung externer Dienstleister

2

Reibungsloser Projektstart mit 20 Silent Bricks innerhalb kürzester Zeit

3

Kontinuierlicher Ausbau der Kapazität

Sicherheit als treibender Faktor

Eine andere wichtige Motivation für die Entscheidung waren die Robustheit und nahezu vollständige Ausfallsicherheit. Immerhin finden viele Transporte zwischen den externen Dienstleistern und dem Bundesarchiv statt, nicht in jedem Fall geht es dabei sanft zu. In jedem Speichermedium wird Erasure Coding eingesetzt, um Datenverlust durch Ausfall zu vermeiden. „Die Gefahr von Datenverlusten ist durch das integrierte Sicherheitskonzept praktisch null. Früher hingegen mussten die Unterlagen bei Ausfällen von USB-Festplatten erneut digitalisiert werden“, sagt Timo Dommermuth. Die Digitalisierung ist nicht immer ganz ohne, besonders, wenn es um die Arbeit mit alten Nitrofilmen geht. Da ist im wahrsten Sinne des Wortes Fingerspitzengefühl gefragt, damit die Filme nicht zerstört werden oder gar Menschen in Gefahr geraten. Die Trägerschicht aus Nitrozellulose, auch Schießbaumwolle genannt, fällt unter das Sprengstoffgesetz. Die Filme müssen deshalb bei einer bestimmten Temperatur gelagert werden, damit es nicht zur Selbstentzündung kommt. Auch die Speicherung der digitalisierten Nitrofilme erfolgt jetzt in geschützten Räumen auf den Silent Bricks. Allein die Kinowochenschauen aus der Zeit des Zweiten Weltkriegs umfassen in digitaler Form schon 20 Terabyte.

Das Speichersystem wächst weiter

Über die Jahre ist das System, das sich flexibel nach dem Baukastenprinzip erweitern lässt, gewachsen. Gerade ist man in einer weiteren Beschaffungsphase. In Kürze sollen

weitere sechs Controller G5000 und rund 30 Silent Bricks mit je 24TB Brutto SSD angeschafft werden. Der Trend geht klar in Richtung Digitalisierung on Demand: Sobald eine Akte aufgrund einer Anfrage angefasst wird, erfolgt gleich deren Digitalisierung in einem Scan-Prozess. Auch deshalb will man die Silent-Brick-Infrastruktur immer weiter ausbauen. Die Experten beim Bundesarchiv gehen davon aus, dass sie ab 2024 rund 60 bis 80 Millionen Seiten pro Jahr digitalisieren können. Bis jetzt liegen 75 Millionen Seiten digital vor – auch im internationalen Vergleich ist das viel. Die Motivation bei den Archivnutzern ist ganz unterschiedlich: Oft sind es Wissenschaftler oder Doktoranden, die Dokumente für ihre Forschung benötigen. Aber auch zur Schicksalsklärung wird immer wieder angefragt: Die Enkelgeneration interessiert sich zunehmend für die Aufarbeitung der Familiengeschichte während des Nationalsozialismus. Doch es gibt auch eher ungewöhnliche Trends: So orientieren sich beispielsweise Spiele-Designer bei der Gestaltung an Plänen von Schlachtschiffen aus dem Zweiten Weltkrieg. Sie alle bekommen ihre Informationen jetzt fast durch die Bank nicht mehr als Papierkopie, sondern digital.

Autorin: Daniela Hoffmann



ERGEBNISSE

1

Geschwindigkeit beim Datentransfer auf bis zu 700 MB / s gesteigert

2

Die Durchlaufzeit, bis Daten im Archiv digital verfügbar sind, mehr als halbiert

3

Keine Datenverluste mehr durch den Transport

Über das Bundesarchiv

Das Bundesarchiv ist eine Bundesoberbehörde und hat den gesetzlichen Auftrag, das Archivgut des Bundes auf Dauer zu sichern und nutzbar zu machen. Aufbewahrt werden Akten, Karten, Bilder, Plakate, Filme und Tonaufzeichnungen in analoger und digitaler Form. Um die 2300 Mitarbeitenden sind derzeit im Bundesarchiv tätig, etwa 100 davon in der IT. Das Archiv wurde 1952 in Koblenz errichtet. Zuvor gab es seit 1919 als zentrales Archiv das Reichsarchiv auf dem Brauhausberg in Potsdam. Die alliierten

Mächte gaben nach Kriegsende beschlagnahmtes Schriftgut an das Bundesarchiv ab. Seit 1955 ist das Bundesarchiv auch für die dauerhafte Sicherung der militärischen Überlieferung des Bundes und seiner Vorläufer verantwortlich. Im Juni 2021 ging zudem die Verantwortung für die Unterlagen des Ministeriums für Staatssicherheit (MfS) der ehemaligen DDR an das Bundesarchiv über.

Partner

Bechtle IT-Systemhaus Bonn/Köln

BECHTLE