



Hilfestellung zur Auswahl eines Speichersystems

Zusammenfassung

Dieses Dokument gibt Überblick zu den verschiedenen Optionen der Datenablage für unterschiedliche Anwendungsfälle in Forschung, Lehre und Administration. Mit der Ermittlung des Bedarfs durch die Hilfestellung zur Speicherauswahl sollte in der überwiegenden Zahl der Fälle eine qualifizierte Entscheidung möglich sein. Dieses Dokument dient zur Übersicht und Verständnisbildung, ersetzt jedoch nicht eine möglicherweise notwendige Beratung.

Die hier vorgestellten Speichersysteme gehören zur Allgemeinversorgung, die durch das Rechenzentrum und Partner Mitgliedern der Universität bereitgestellt werden. Für höhere Bedarfe in Forschungsprojekten können Betriebs- und Beteiligungsmodelle vorgeschlagen werden, bei denen Beschaffungen innerhalb solcher Projekte in Form von Compute oder Speicher in die Basisinfrastruktur der Maschinsäle integriert werden. Diese Projekte erhalten sich im Rahmen der Benutzungsordnungen des Rechenzentrums Souveränität in der Nutzung ihrer eigenen Ressourcen und entlasten sich von allgemeinen betrieblichen Anforderungen wie der laufenden Administration und Aktualisierung der Systeme.

Persönliches Home-Verzeichnis und Speicher für Arbeitsgruppen

Home-Verzeichnisse werden automatisch mit dem Erstellen eines Uni-Accounts angelegt und bieten anfänglich Platz für Daten bis zu einem Gesamtvolumen von 20 GB (bei Bedarf kann diese Quota erhöht werden). Arbeitsgruppen und Professuren erhalten kostenfrei auf Anfrage einen Speicherbereich in Höhe von 250 GB, welcher bei Bedarf gemäß der Ser-

Albert-Ludwigs-Universität
Freiburg

Rechenzentrum

Allgemeine Verwaltung

Jan Leendertse
Research Data Management
Group

Hermann-Herder-Str. 10
79104 Freiburg

Tel. 0761/203-4646

fdm@mail.uni-freiburg.de
<https://rdmg.uni-freiburg.de>

Freiburg, 27.09.21

■ vicebeschreibung erweitert werden kann. Der Speicherbereich für Arbeitsgruppen / Professuren wird meistens in Form eines sogenannten Gruppenlaufwerks realisiert, welches eine granulare Rechteverwaltung ermöglicht. Falls eine granulare Rechteverwaltung nicht notwendig ist, kann auch über eine einzelne Freigabe beantragt werden. Alternativ kann neben dem Gruppenlaufwerk ein NFS Export (für Linux geeignet, wobei die Zugriffssteuerung über IP-Adressen basiert) bezogen werden. Das Speichersystem ist aus der Produktreihe Isilon der Firma Dell/EMC, welche in den Serverräumen MS II (Rechenzentrum) und MS III (KG 2) (georedundant) betrieben werden. In diesen Verzeichnissen werden regelmäßig automatisch Snapshots erstellt (Zustand einer Datei zu diesem Zeitpunkt) und regelmäßig zentrale Backups im TSM (Bandlaufwerke) geschrieben.

Aufgrund der redundanten Speicherung sowie der Snapshots, werden die Daten in diesem Speichersystem sicher gespeichert und bieten die Möglichkeit, versehentlich gelöschte Dateien sowie verschiedene Versionen einer Datei selbständig wiederherzustellen. Da Snapshots nur lesbar bereitgestellt werden, sind sie vor den typischen Trojaner-Angriffen geschützt.

Der Zugriff auf ein Gruppenlaufwerk setzt zwingend einen Account als Mitglied der Universität Freiburg voraus.

Anfragen zu diesem Speichersystem können an fileservers@rz.uni-freiburg.de (Ticket-Queue-Frontend zum zentralen Fileserver) gerichtet werden.

Wissenschaftliches Arbeitsverzeichnis und Speicher für wissenschaftliche Workflows

Daten im wissenschaftlichen Zusammenhang unterliegen meistens bestimmten Anforderungen, die beispielsweise durch Fördergeber oder technische Gegebenheiten (schnelles Messgerät, welches mit hoher Datenrate schreibt) definiert werden. Solche Systeme sind potenziell fach- bzw. gruppenspezifisch. Als übergreifendes System kommt für diese Zwecke bwSFS (Storage-for-Science) in Frage. Wichtige begleitende Daten aus der Verwaltung des Projekts oder des Vorhabens (Publikationen, Reisekostenabrechnungen, ...) sollten hier nicht abgelegt werden sondern im vorgenannten "Persönliches Home-Verzeichnis und Speicher für Arbeitsgruppen".

Wissenschaftliche Arbeitsverzeichnisse dienen primär der Forschung (insbesondere, wenn sie DFG-gefördert sind). Sie sind potenziell für automatisierte

- Workflows im Lese-/Schreibzugriff verfügbar: Auf diese Weise können beispielsweise Instrumente oder Compute-Systeme Daten ohne direkte Nutzerinteraktion schreiben oder automatisch ablaufende Hintergrunddienste Zugriff auf die wissenschaftliche Daten erhalten. Diese Hintergrunddienste könnten unter anderem Metadaten aus dem wissenschaftlichen Arbeitsverzeichnis beschaffen und automatisch an Datensätze anhängen, die sie an anderer Stelle ablegen (z.B. automatisches Verschieben von Daten aus Mikroskopen, Spektrometern etc. in das wissenschaftliche Arbeits- oder Gemeinschaftsverzeichnis). Gemeinschaftsverzeichnisse können zum Teilen von Daten mit Kolleg*innen innerhalb einer Arbeitsgruppe eingerichtet werden

Anfragen zu diesem Speichertyp können allgemein an die Research Data Management Group unter fdm@mail.uni-freiburg.de gestellt werden, da sie vielfach im Zusammenhang mit dem Forschungsdatenmanagement stehen.

Object Storage

Im Rahmen des bwSFS wird eine erhebliche Menge an Speicherplatz in Form von Object Storage für "lauwarme" und kalte Daten zur Verfügung gestellt. Speziell handelt es sich dabei um ein Speichersystem, welches weitgehend kompatibel mit Amazon S3 ist. Dieses Interface bietet sich insbesondere für das Bereitstellen von Daten auch über Grenzen eines lokalen Campus-Netzwerkes an. Einfache Formen wären beispielsweise das im Vergleich zum TSM-Backup schnellere Weg- und Rücksichern von Datenpaketen oder -containern aus dem parallelen Filesystem.

Der Speicher ist für die kalten Daten als Speicher-Backend für Datenrepositorien, wissenschaftliche Workflow-Systeme und andere Anwendungen geeignet, bei denen Verfügbarkeit und Erhaltung von Daten im Vordergrund stehen. Dieser Speichertyp wird daher in bwSFS für das Forschungsdatenmanagement eine wesentliche Rolle spielen. Für den direkten Zugriff durch Endnutzer und für Anwendungsfälle, bei denen sich der Inhalt einzelner Dateien oder Objekten häufig ändert, ist dieser Speicher ungeeignet. In diesen Fällen sollte ein Dateisystem-Speicher verwendet werden. Anfragen zu diesem Speichertyp können allgemein an fdm@mail.uni-freiburg.de gestellt werden, da sie vielfach im Zusammenhang mit dem Forschungsdatenmanagement stehen.

— Storage in der bwCloud

Hierbei handelt es sich um speziellen Speicherplatz, der mit in den bwCloud angelegten virtuellen Maschinen bereitgestellt wird. Einerseits besteht die Option mit dem Rootfilesystem eine bestimmte Größe zu erhalten, die durch die für die virtuelle Maschine gewählte Ausstattung bestimmt wird. Andererseits ist es auch möglich, Volumes als weitere virtuelle Festplatte einzubinden. Diese können auch durchaus größer werden.

Die bwCloud macht systembedingt wenig Verfügbarkeitszusagen, so dass wichtige Daten nicht hier, sondern in anderen geeigneten Speicherbereichen (persönliches/wissenschaftliches Home) abgelegt werden sollten. Sie ist für automatisch angelegte und wieder zu löschende Maschinen optimiert (und in der zugesicherten Zuverlässigkeit nicht mit dem angebotenen VMware ESX vergleichbar).

Storage für die VMware ESX Virtualisierung

Dieses Speichersystem ist vergleichbar mit dem System, das unter "Storage in der bwCloud" dargestellt wurde. Es handelt sich um einen hochperformanten Speicher, der speziell auf die Virtualisierungstechnologie von VMware ESX Enterprise abgestimmt ist. Aktuell werden hierfür Systeme der Firma Tintri verwendet, die sich durch eine besonders hohe Integration mit VMware ESX auszeichnen. Auf diesem Speicher liegen prinzipiell nur die Betriebssysteme der jeweiligen virtuellen Maschinen (VMs). Sollten Dienste, die auf der VMware Virtualisierungsinfrastruktur betrieben werden, größere Datenmengen benötigen, dann können diese über NFS- oder SMB-Mounts von dem zentralen Isilon-Speichersystem eingebunden werden. Es gibt aber Sonderfälle, in denen das Einbinden von externem Speicher über NFS/SMB zu Problemen führen kann (z.B. hochfrequentierte Datenbanksysteme). In diesem Fall kann nach Absprache auch Datenvolumen auf dem Virtualisierungsspeicher vorgehalten werden. Es kann in solchen Fällen Speicher vom zentralen DELL/Isilon Speicher per NFS oder SMB eingebunden werden. Dieser Speicher ist technisch dasselbe System, das auch die persönlichen Home-Verzeichnisse und die Gruppenlaufwerke vorhält.

Durch die hohe Integration des Systems mit VMware ESX werden Redundanzen bzw. Backups über die Virtualisierungsebene über Snapshots gesteuert.

- Anfragen zur Virtualisierung und diesem Speichertyp können an virtualisierung@rz.uni-freiburg.de gestellt werden.

Sync & Share

bwSync and Share wird als Landesdienst durch das KIT bereitgestellt und basiert technisch auf der NextCloud-Software, welche neben geteiltem Speicherplatz auch die Möglichkeit mitbringt, Textdateien, Tabellenkalkulationen etc. gemeinsam zu bearbeiten. Es sind weitere Kollaborationsmodule eingebaut. Die Universität Freiburg beteiligt sich an der Finanzierung. Es umfasst in der Voreinstellung 50 GByte pro Nutzer. Im Gegensatz zu den anderen Speicher-Optionen der Universität Freiburg steht dieser Speicherbereich auch ohne Umwege eines VPNs weltweit zur Verfügung. Durch die Bereitstellung von Apps für iOS und Android kann dieser Speicher auch von mobilen Geräten aus verwendet werden. Der First-Level-Support ist über die Ticket-Queue bwsyncandshare@rz.uni-freiburg.de erreichbar.

Spezialspeicher: High Performance Parallel Filesystem

Solche Dateisysteme finden sich typischerweise in HPC-Umgebungen, in denen beispielsweise massiv parallele Jobs ausgeführt werden. Hierzu werden die notwendigen Dateien üblicherweise aus anderen Speicherbereichen auf das parallele Filesystem kopiert, um dann zeitnah verarbeitet zu werden. Nach der Verarbeitung sollten die Ergebnisse auf andere Speicherbereiche, z.B. das wissenschaftliche Arbeitsverzeichnis oder in den Object Storage von bwSFS, gesichert werden. Da der Speicherplatz auf diesen Systemen von vielen Gruppen und Forschenden geteilt wird, arbeitet man häufig mit zeitlich begrenzten Reservierungen. Nach Ablauf solcher Reservierungen werden Speicherbereiche automatisch durch das System bereinigt.

Solche parallelen Dateisysteme sind oft an spezielle Infrastrukturen, wie beispielsweise Infiniband oder OmniPath oder sehr schnelles Ethernet, angebunden und stehen daher selten außerhalb eines HPC-Clusters zur Verfügung. Weiterhin bieten sie oft keine Rechteseparierung, so dass alle Bereiche eines eingebundenen Systems vom Administrator gelesen und geschrieben werden können. Solche Filesysteme arbeiten lediglich mit der Redundanz, die durch Hardware-RAID angeboten wird, da sonst die angestrebte

Performance nicht erreicht wird. Techniken, wie Snapshots werden ebenfalls nicht angeboten, so dass kein Zugriff auf frühere Versionen oder versehentlich gelöschten Daten möglich ist.

Wegen der großen Datenmengen auf solchen Systemen findet kein Backup statt, so dass in einem Disaster-Fall Daten unwiederbringlich verloren gehen können. Hierzu können regelmäßige Sicherungen in den Object Storage genutzt werden.

Die Ticket-Queue für den NEMO Support (welche auch das parallele Filesystem umfasst) ist unter enm-support@hpc.uni-freiburg.de zu erreichen. Unter hpc-support@hpc.uni-freiburg.de ist die Ticket-Queue für den generellen HPC-Support erreichbar (andere Cluster im Land und deren jeweilige parallele Filesysteme).

Sichere Dokumentenablage (Groups/BSCW)

Für den gruppeninternen Austausch von sensiblen Daten bietet sich der Arbeitsgruppenserver (BSCW) der Universität Freiburg an. Im Gegensatz zu den anderen Speicherangeboten der Universität kann dieser Speicherbereich nicht mit einem Laufwerksbuchstaben eingebunden werden. Die Verwaltung der Daten erfolgt über das angebotene Webportal. Eine Einbindung über das WebDAV-Protokoll ist möglich.

Die Ticket-Queue für den Arbeitsgruppenserver ist unter der Adresse groupserver@rz.uni-freiburg.de zu erreichen.

Datenverteilung und Ablage für die Lehre

Wenn Unterlagen zu Lehrveranstaltungen in verschiedenen Formen verteilt werden sollen, bietet sich hierzu das Learning-Management-System Ilias an. Dieses ist bei der überwiegenden Zahl der Studierenden bekannt und wird in vielen Veranstaltungen verwendet. Der Zugriff auf die Materialien kann dabei auf jeweilige Gruppe eingeschränkt werden.

Die Ticket-Queue für die Lernplattform ILIAS lautet ilias@rz.uni-freiburg.de