



# Datensicherung im Wandel

Thomas Eifert

# RWTH aus der Perspektive des zentralen Backup-/Restore-Service

---

Fast 100%-Abdeckung der Einrichtungen als Nutzende des zentralen Backup/Restore Service

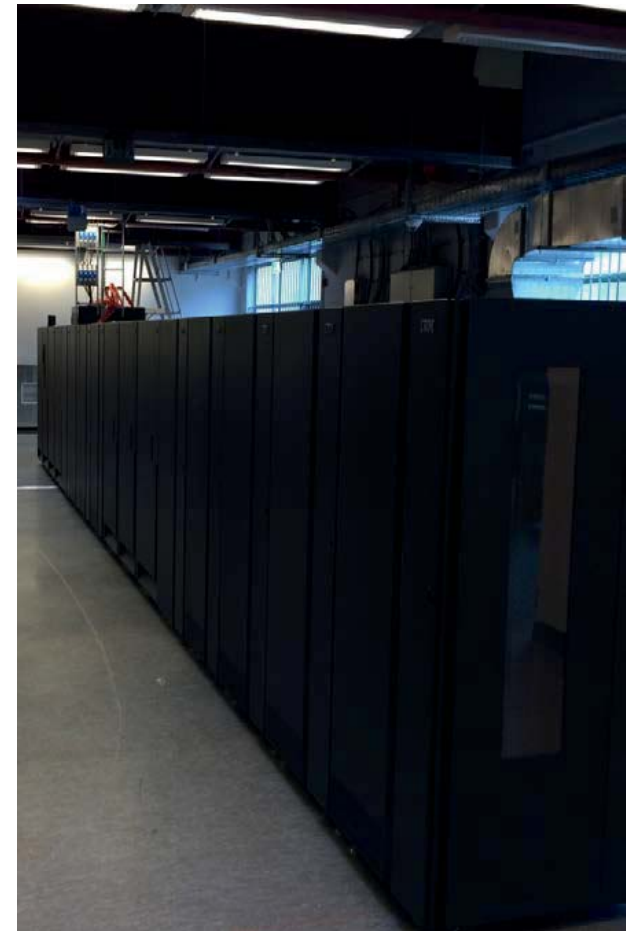
Aktuelle Infrastruktur für TSM (seit Juli 2015)

- 4 pSeries S822, jew. 1 TB RAM
- 1 PB Disks für StgPools
- TS 3500 Library
- 48 x IBM TS1150 tape drives

Seit 2010: Service Level Agreement mit Einrichtungen:  
„any restore at any time with client's wirespeed“

presented on TSM Symposium 2011 and in PIK publication,  
Vol. 35(3), p. 195–198, DOI: [10.1515/pik-2012-0032](https://doi.org/10.1515/pik-2012-0032)

Service für Uni Paderborn und FH Aachen



# Datensicherung ?

## Definition:

Durch technisches Versagen, versehentliches Löschen oder durch Manipulation können gespeicherte Daten unbrauchbar werden bzw. verloren gehen. Eine Datensicherung soll gewährleisten, dass durch einen redundanten Datenbestand der IT-Betrieb kurzfristig wiederaufgenommen werden kann, wenn Teile des operativen Datenbestandes verloren gehen.

aus: BSI IT Grundschutz

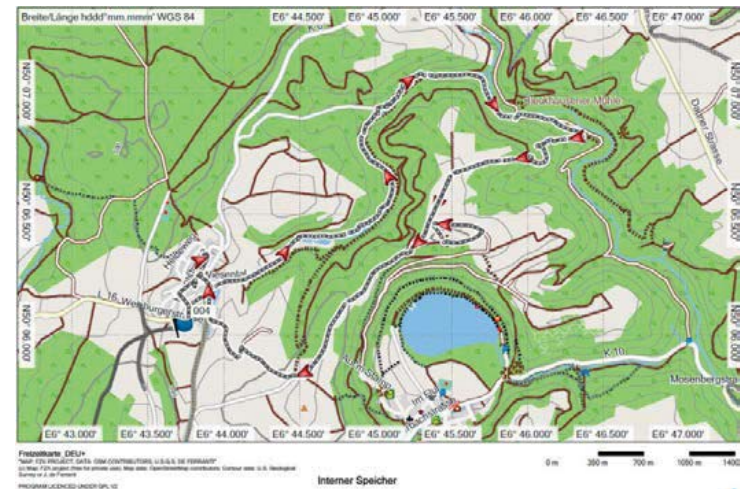
## Kürzer:

Schutz von Datenbeständen gegen Verlust

Also: „Es kommt nichts weg“

→ Ziel klar !

Wege zum Ziel ⇔ Ausgangspunkte



# „Datensicherung“

---

- Datensicherung: Gesamtheit der Maßnahmen zum Schutz gegen Datenverlust
- Backup: (Blindes) Abschreiben eines Datenbestandes auf separate Plattform an separatem Standort
- Tape: Speichermedium

# Datenhaltung: Was haben wir?

---

- Online-Storage („General purpose“)
  - Instituts-Fileserver (~10 PB)
  - Lokale Platte auf Endgerät
- Online-Storage
  - HPC (~1.5 PB)  
Singuläre Größen
  - Kollaborations-Plattformen
    - Web-Server, Wikis,
- Nearline
  - Archiv

➔ Speicherorte wertvoller Daten ➔ vor Verlust zu sichern

## Was haben wir noch:

---

- Transaktions-Speicher („strukturierte Daten“)
  - Datenbanken
    - Verwaltung
      - ERP (~100 TB)
      - Campus Management (~0.6 TB)
    - IT-Administration
      - IdM
      - NOC
      -
    - Datenbanken für Nutzer-Dienste
    - FD-Metadaten
  - Sharepoint

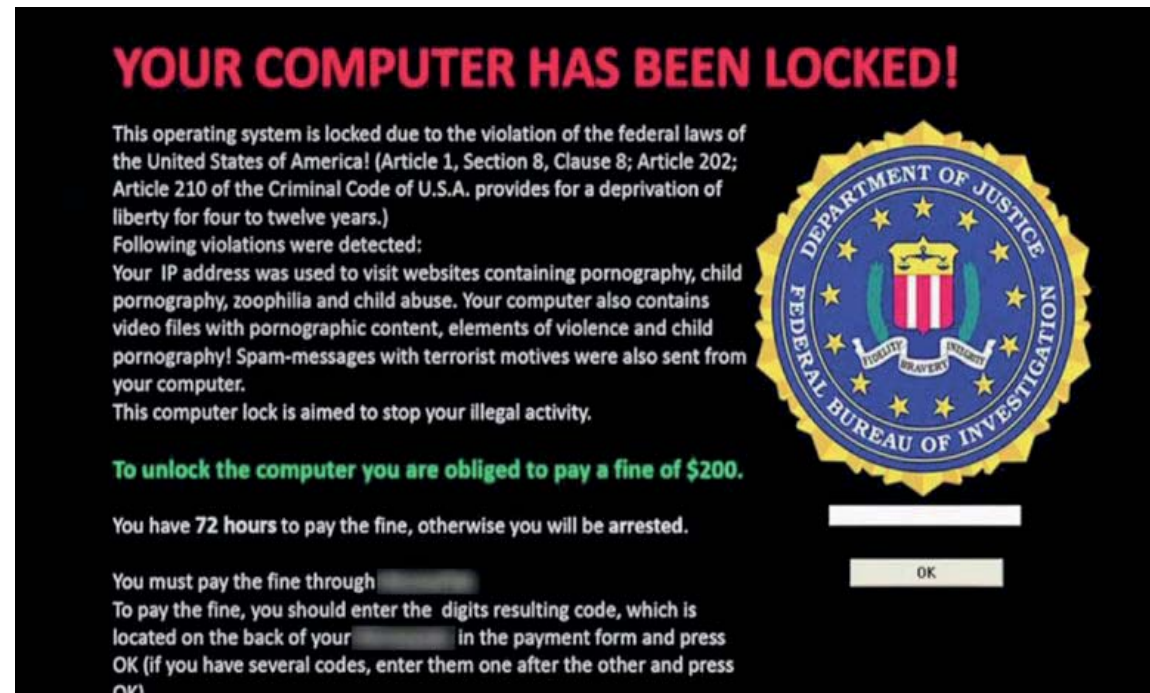


Benötigt Sicherung  
auf  
Anwendungsebene

# Datensicherung: Schutz von Datenbeständen gegen Verlust

Gegen Fehlbedienung / Software-Fehler auf Anwendungsebene

- Kopieren → Backup
- Versionierung, ggf. mit Überschreib- / Lösch-Schutz
- Verzögertes Ausführen von Löschungen



# Datensicherung: Schutz von Datenbeständen gegen Verlust

Gegen Ausfall des Speichersystems /-Mediums (Hardware, Gebäude, ..)

- Kopieren
  - Replikation (Cloud),
  - Backup
  - Bei großen Speichersystemen: Restore kaum in sinnvoller Zeit möglich
- Verteilen → Redundanz, EC



Foto: © Ralf Röger



# Datensicherung: Schutz von Datenbeständen gegen Verlust

Gegen Software-Fehler auf  
Speichersystem-Ebene

→ Kopieren mit Plattform-Wechsel

→ Backup

**Laut einer Studie an einer  
englischen Universität ist es  
egal, in welcher Reihenfolge  
die Buchstaben in einem  
Wort sind.**



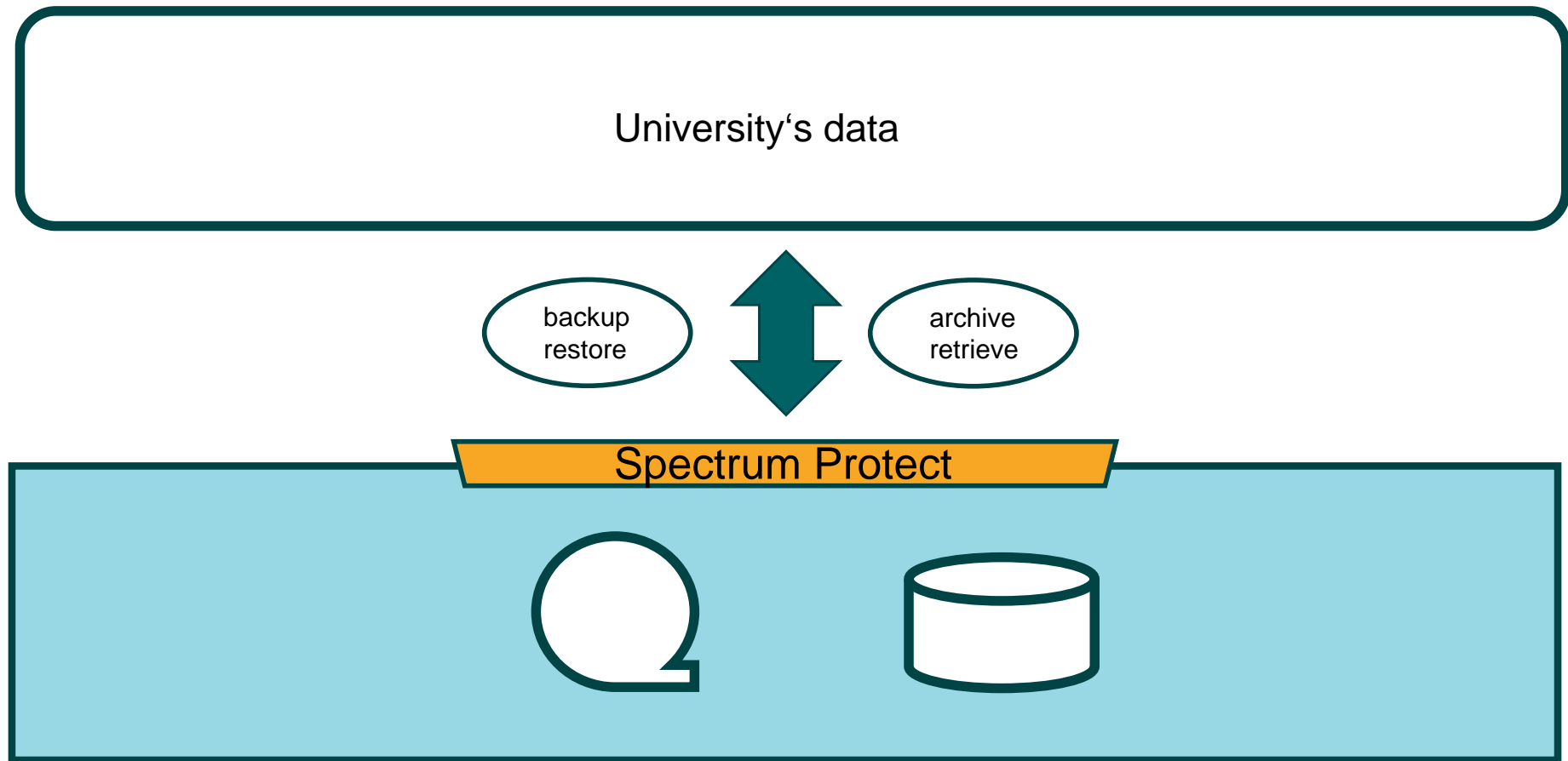
**Luat enier sidtue an eienr  
elgnhcsien uvrnäiett, ist es  
eagl in wcheler rhnfgeeloie  
die bstuchbaen in eniem  
wrot snid.**

# Sicherungsszenarien

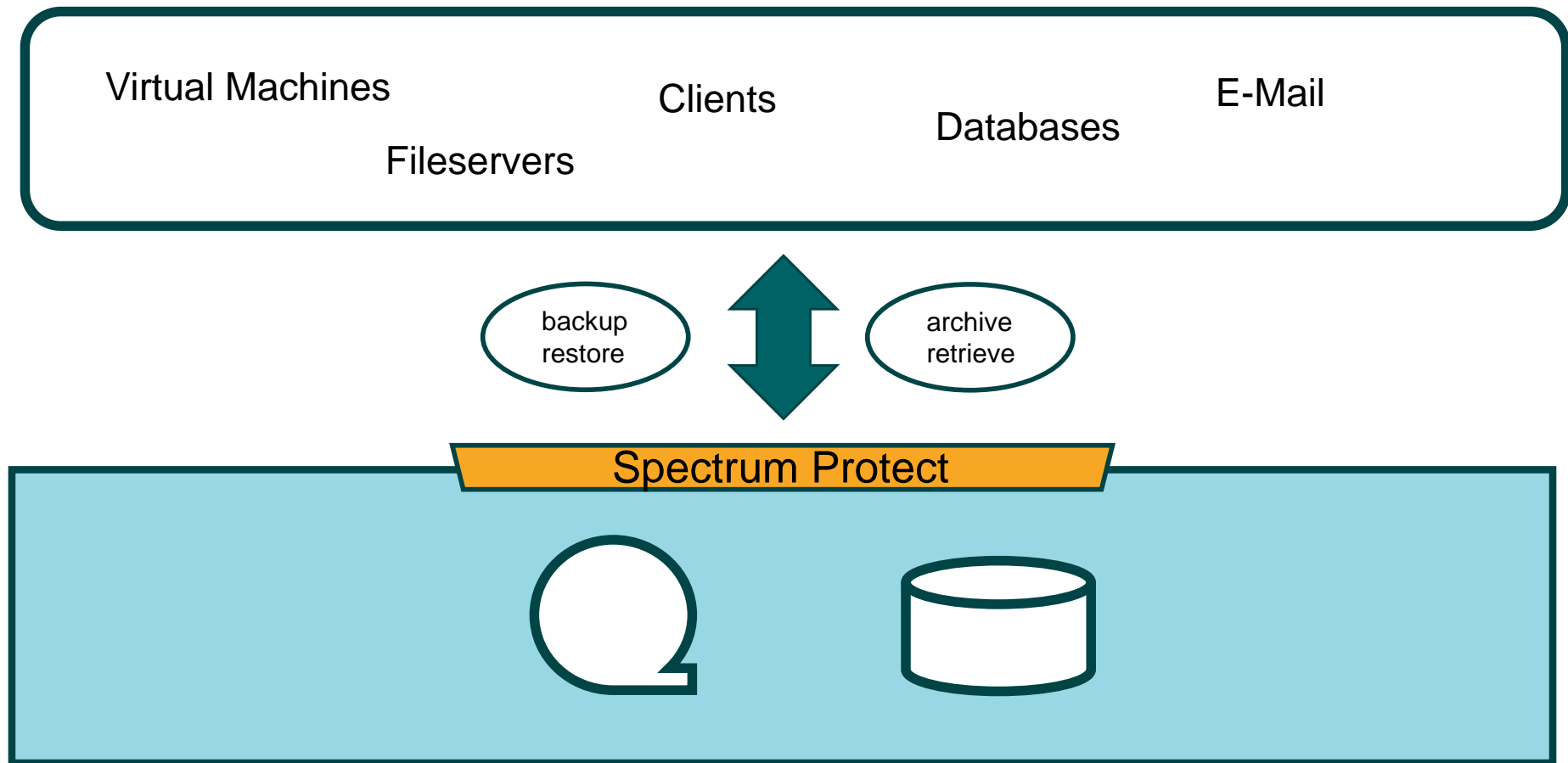
---

- Was wird wo gespeichert?
- Wie gesichert?
- Und wie könnte das künftig aussehen?

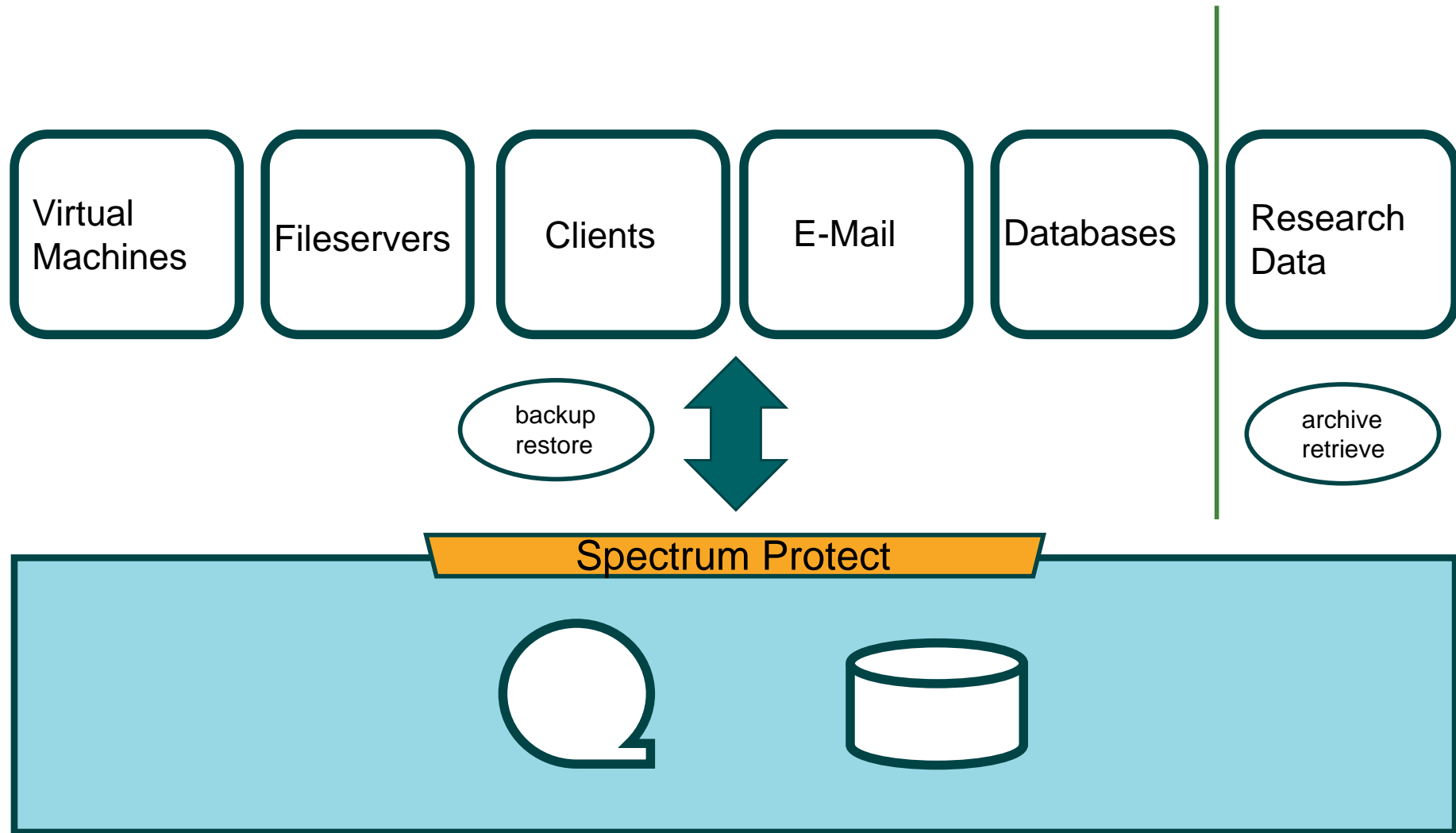
# Backup aller Daten



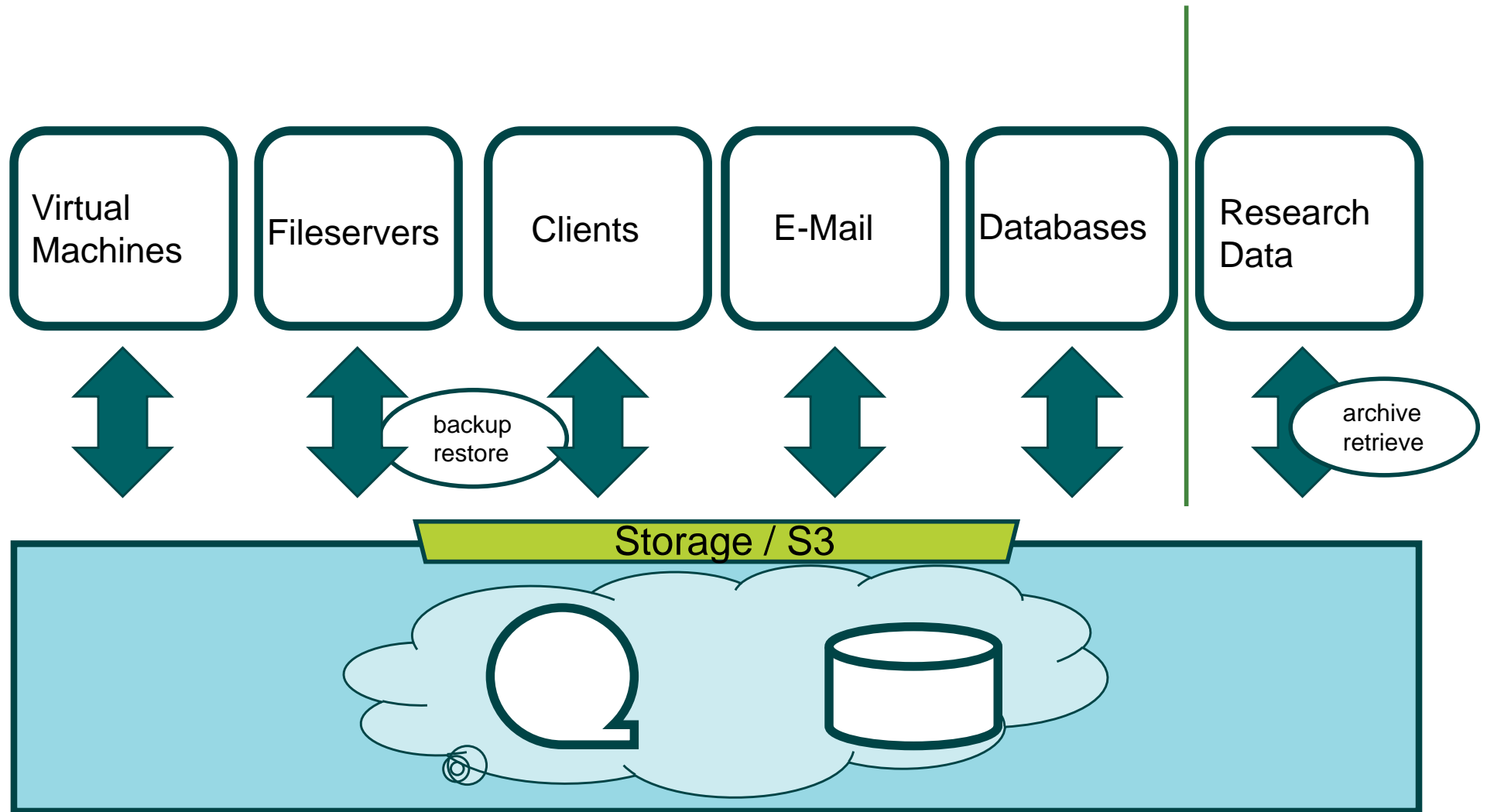
# Backup: Differenzierte Datenquellen



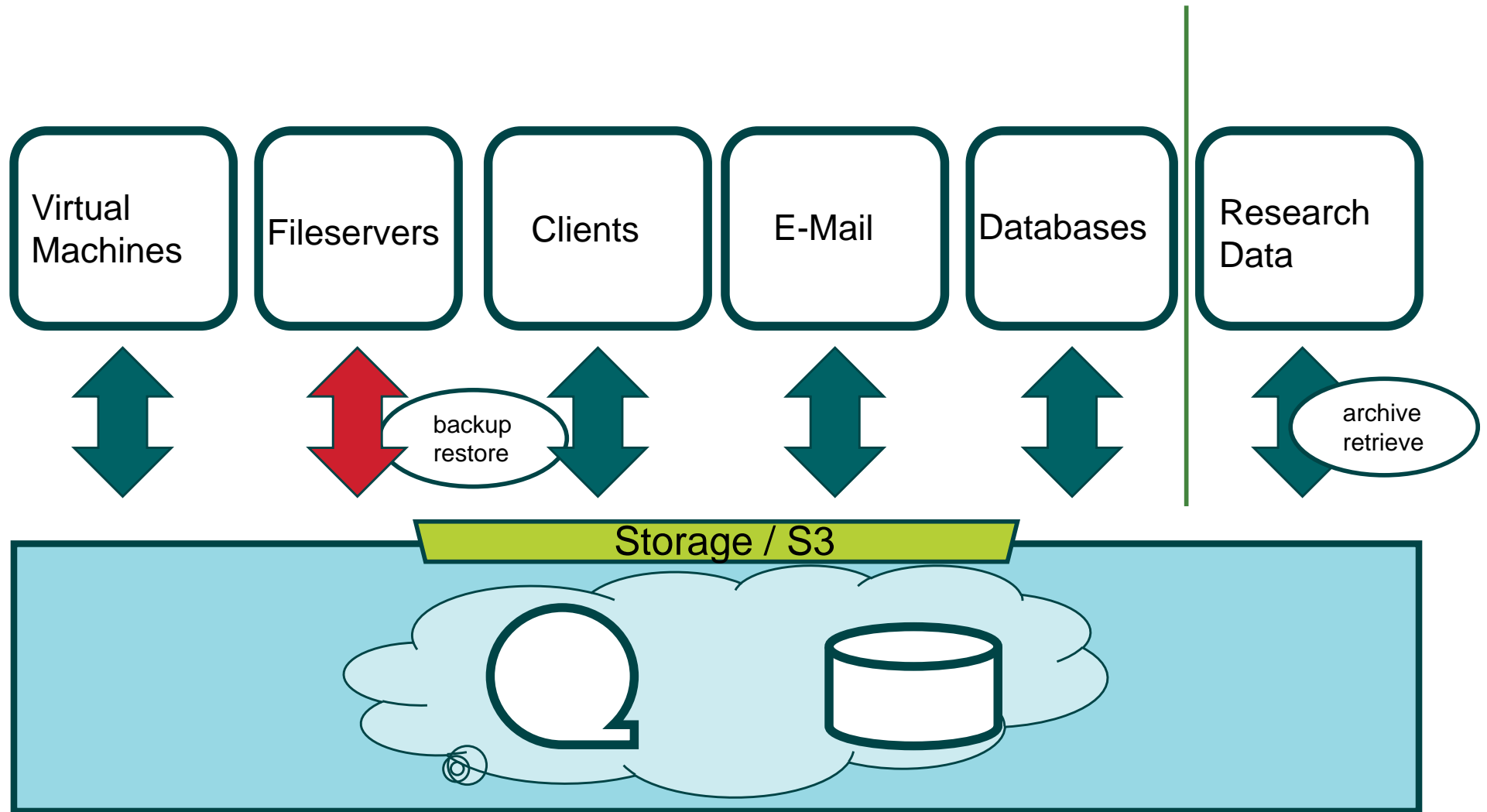
# Backup: Einer für alle?



# Backup: .. Oder für jeden das Beste?



# Dienste-spezifische Sicherungsstrategie



# „Datensicherung“

---

- Datensicherung: Gesamtheit der Maßnahmen zum Schutz gegen Datenverlust
- Backup: (Blindes) Abschreiben eines Datenbestandes auf separate Plattform an separatem Standort
- Tape: Speichermedium
- Archiv: Bewusstes Ablegen eines (abgeschlossenen) Datenbestandes auf zuverlässigem Langzeitspeicher



# Sicht auf Daten

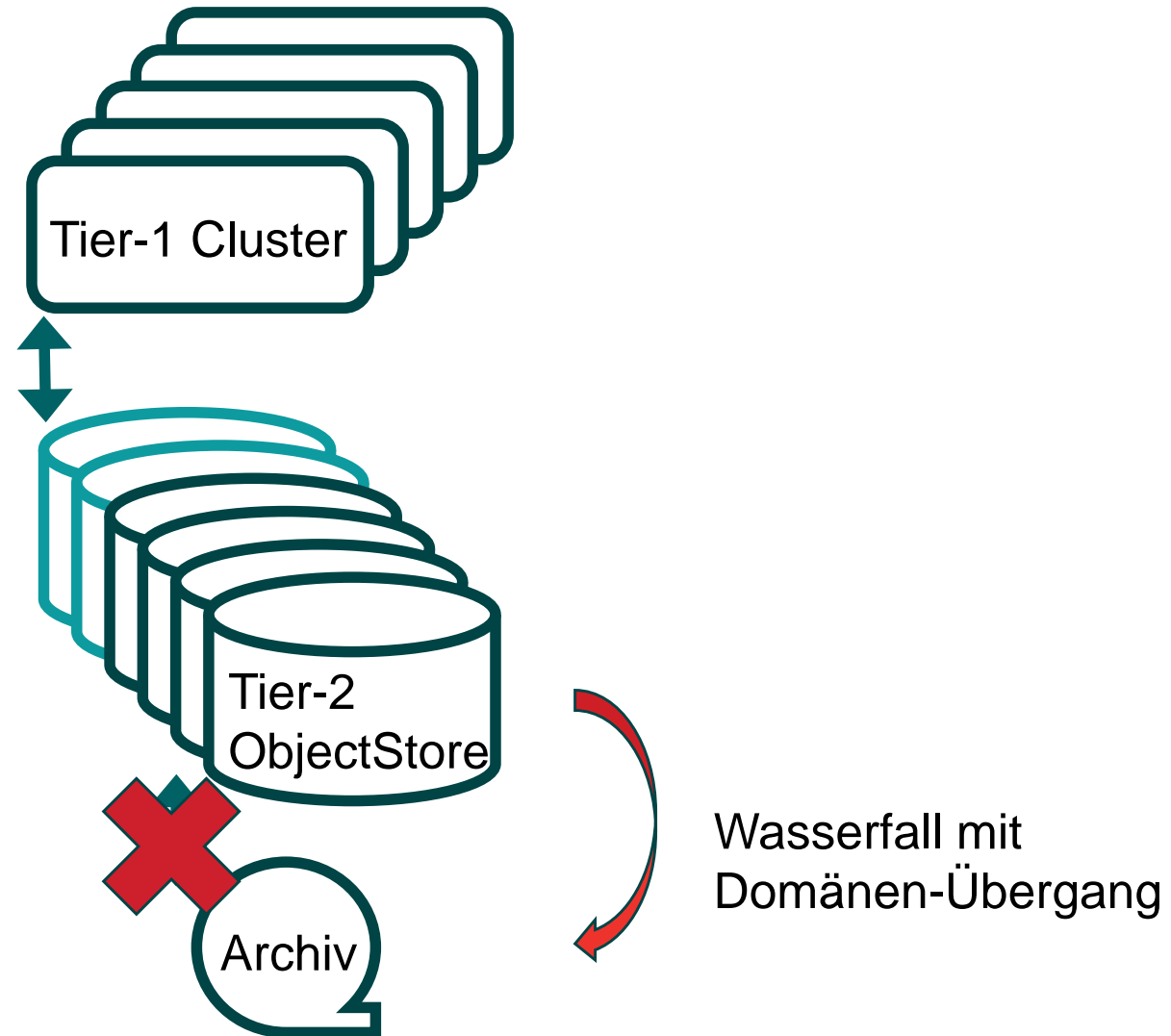
---

- Traditionelle Sicht auf Daten:
    - Die wertvollsten Daten sind die aktuellen
    - Die Relevanz nimmt mit der Zeit ab
    - Annahme (z.B. in Backup-Community):  
archivierte Daten müssen verwahrt werden, in der Praxis wird  
nie mehr drauf zugegriffen

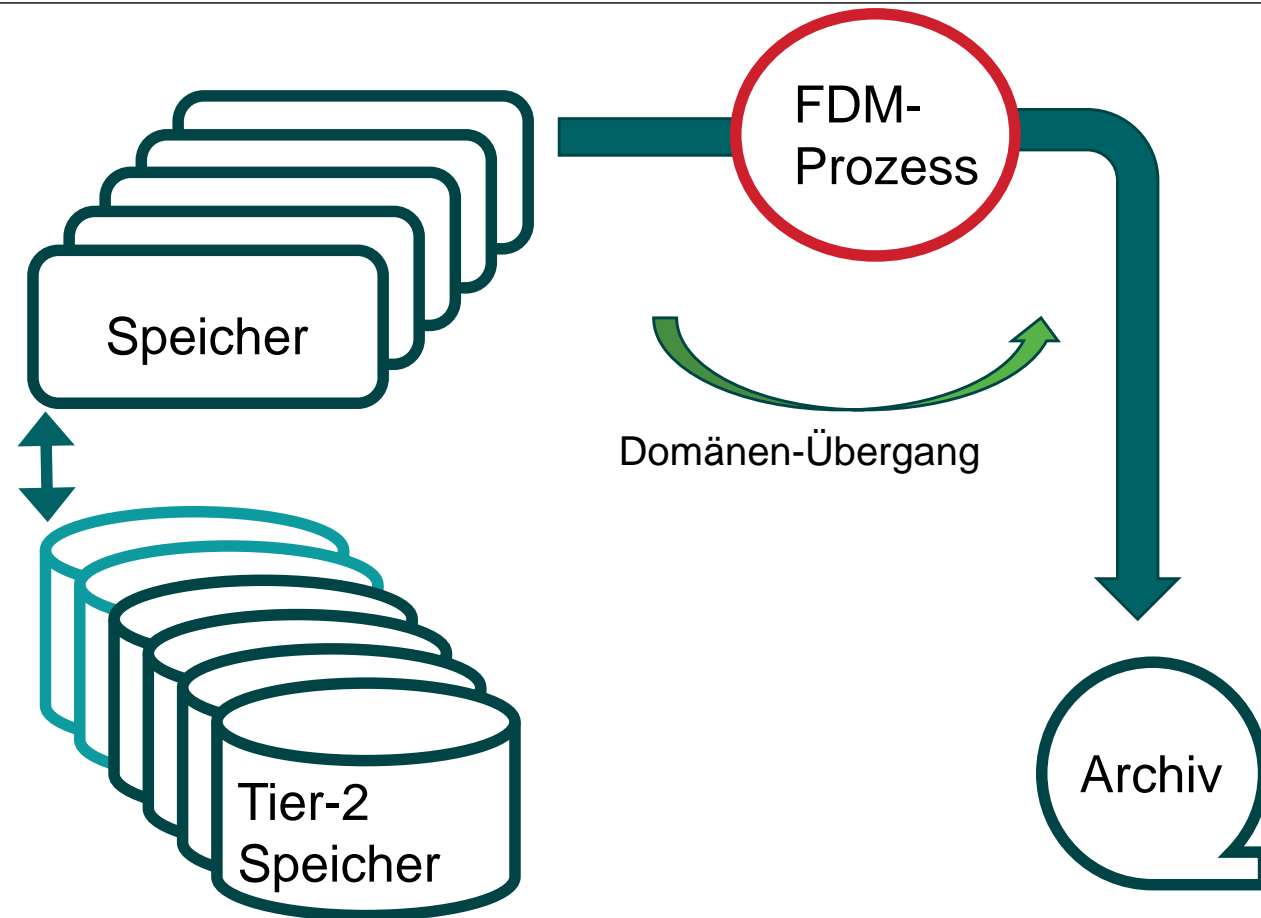
→ Speicher-Hierarchie getriebenes Wasserfall-Modell
  
  - FD-Sicht auf Daten:
    - Es werden viele Experimente gemacht
    - Im Zuge der Auswertung wird sortiert und angereichert
    - Daten, die zu Erkenntnissen führen oder Thesen untermauern,  
sind die wertvollsten und müssen archiviert werden.

→ Das Archiv enthält das „Daten-Gold“
- Widerspruch

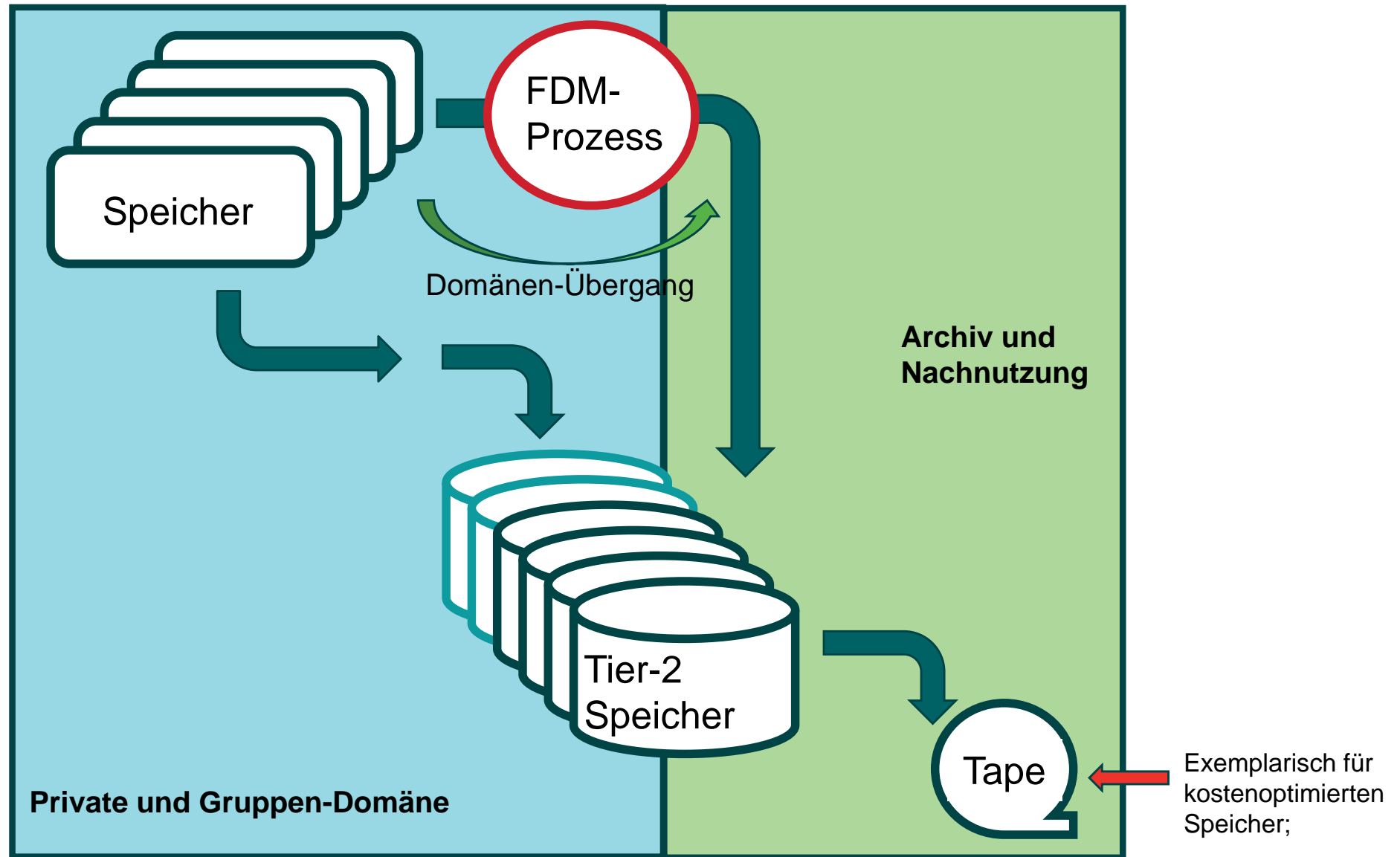
# FD-Storage: 3-Tier-Speicherhierarchie (Wasserfall-Modell)



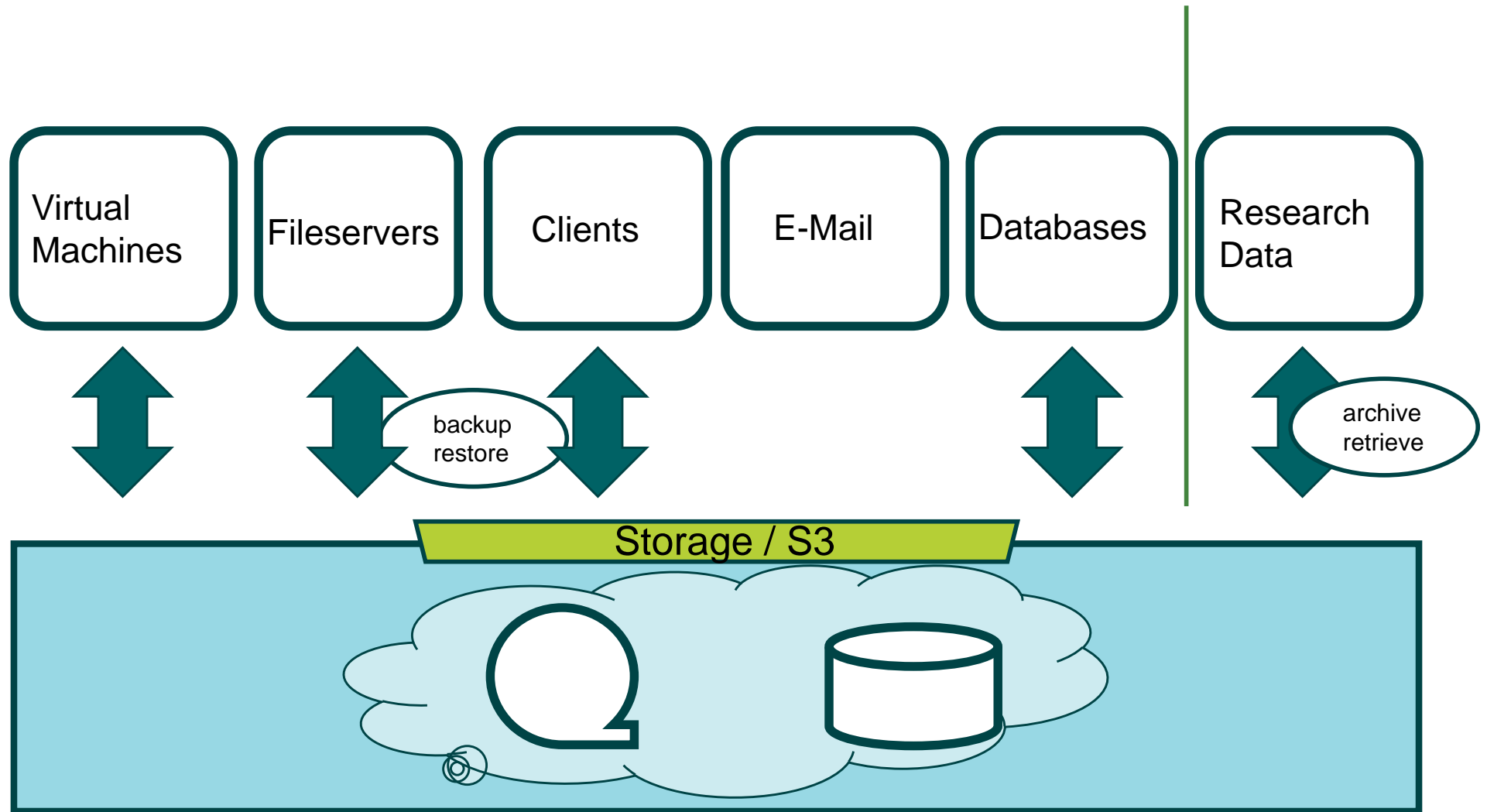
## FD-Storage: 3-Tier-Speicherhierarchie (Archiv auf Tape)



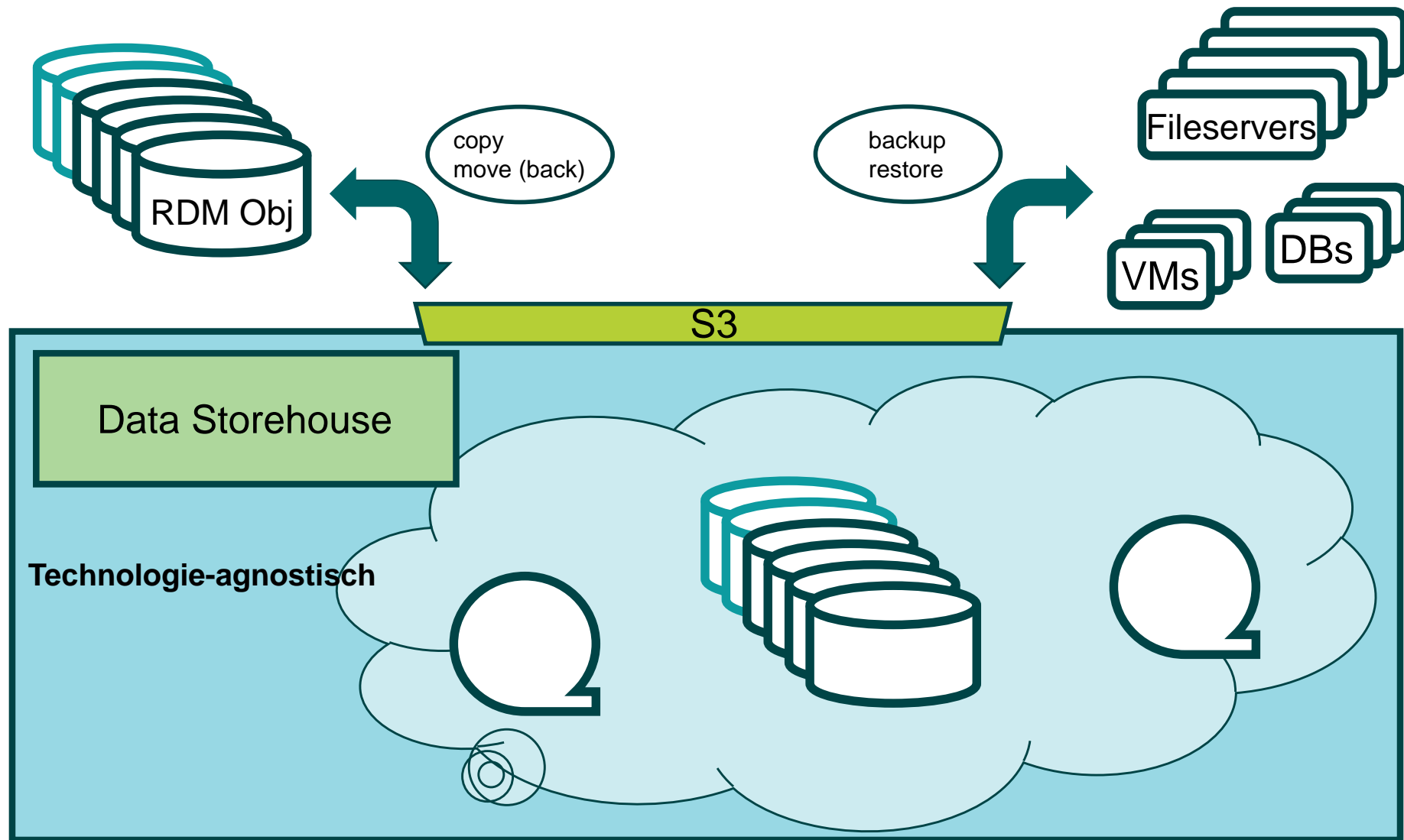
# FD-Storage: 3-Tier-Speicherhierarchie (Archiv mit Speicherhierarchie)



# Data Storehouse: hochvolumiger sicherer Speicher



# Data Storehouse: hochvolumiger sicherer Speicher



# Zusammenfassung

---

- **Daten als Grundlage von Wissenschaft**  
→ **Sichere Datenhaltung wird immer wichtiger**
- **Objekt-Speicher: Abstraktion zwischen Anwendungen und Speicherhierarchie**
- **Differenzierter Einsatz von Speichermedien und -Methoden**

