



IBM TotalStorage™

# Langzeitarchivierung von Massendaten mittels Information Lifecycle Management

**Hartmut Grund**

**IBM Stuttgart**

# Agenda

- **Archivierung**
- **Information Lifecycle Management**
- **IBM TotalStorage Data Retention 550 Solution**
- **Lösungsszenarien**

# Archivierungslösungen

## Warum ist Archivierung (plötzlich wieder) ein Thema in der IT ?

### Gründe:

- **Exorbitantes Datenwachstum**
  - ▶ **Durch neue Anwendungen und Möglichkeiten (Bilder, Video, Musik, File-Attachments)**
  - ▶ **Allgemeiner email-Verkehr**
  - ▶ **Generelles Jäger- und Sammler Syndrom**
  
- **Veränderte Geschäftsbeziehung zwischen Firmen und Kunden**
  - ▶ **Geschäftsbeziehungen, auch vertraglich relevante, werden per email abgewickelt**
  
- **Gesetzliche Vorgaben**
  
- **Archivierungstechnologien unflexibel oder veraltet**
  
- **Kostensituation**

## Archivierung / Data Retention

### Definitionen und Anforderungen

#### **Was ist Archivierung ?**

- ▶ **langfristige, revisions sichere Aufbewahrung von Dokumenten**
- ▶ **Archivierung der Dokumente, wenn sie für den unmittelbaren Geschäftsbetrieb nicht mehr benötigt werden**

#### **Was ist Data Retention ?**

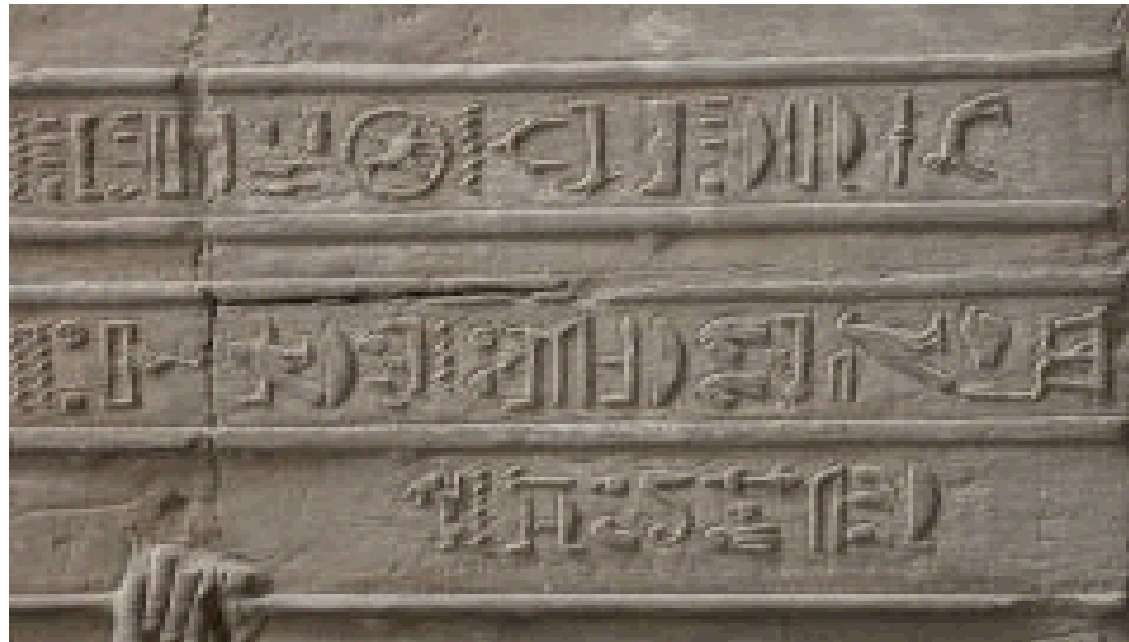
- ▶ **Data Retention befasst sich mit**

- **dem Erhalt,**
- **dem Schutz und den**
- **Aufbewahrungsregularien**

**von archivierten Daten**

## Anforderung an die Archivierung

- **Verfahren der Datenspeicherung** soll sicherstellen, daß die Daten
  - ▶ nicht gelöscht oder
  - ▶ verändert werden können
  - ▶ entsprechend der vorgeschriebenen Zeit aufbewahrt werden
  - ▶ in entsprechender Form wieder gelesen werden können
  - ▶ ggf. eine Kopie erstellt wird
  - ▶ Daten müssen jederzeit les- und auswertbar sein

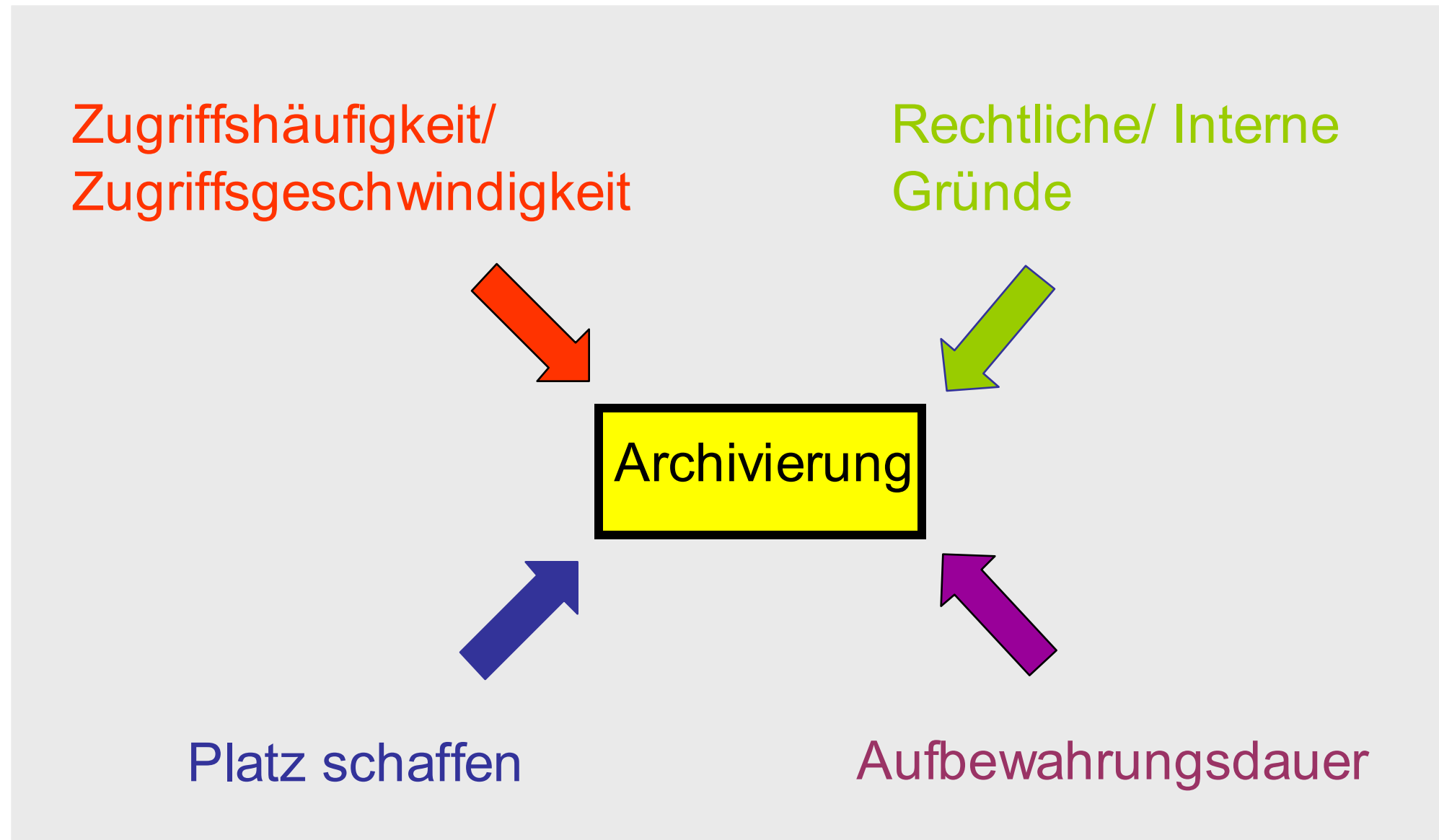


## Anforderung an die Archivierung

- **Technologiewandel berücksichtigen**
  - ▶ Lebenszyclus von IT-Systemen 5-7 Jahre
  - ▶ Anwendungen und Datenformate ca. 10 Jahre
  - ▶ Magnetplatten werden gewöhnlich alle 3 – 5 Jahre ausgetauscht
  - ▶ Magnetbänder halten 10 Jahre
- **Migrationsaufwand** in neue Technologien/Anwendungen einbeziehen
- **Kostenbetrachtung** als ein Entscheidungskriterium
  - ▶ Worauf soll grundsätzlich archiviert werden (Tape, Disk, Optical)
- **Archivierung analog der Wichtigkeit von Daten (ILM)**



## Archivierung Anforderungsprofile





## Istsituation heutiger Archivierungslösungen

- **Archivierungslösungen basieren zum großen Teil auf WORM Technologie (Optische Platten)**
- **WORM garantieren die Nicht-Veränderbarkeit der gespeicherten Daten (WORM=Write Once, Read Multiple)**
- **Optische Platten werden von vielen Unternehmen preferiert aus internen Revisionsgründen oder gesetzlichen Vorgaben**

### **Alternativen ?**

**-Worm-Tape**

**-Soft-WORM mittels Software**

## Auszug aus dem Grundschutzhandbuch des BSI

### Ermittlung der rechtlichen Einflussfaktoren für die elektronische Archivierung

Im Folgenden werden einige Quellen genannt, die in Deutschland typischerweise zu berücksichtigen sind:

- **Bürgerliches Gesetzbuch (BGB)**  
Hier werden insbesondere Anforderungen an die Rechtsgültigkeit von Dokumenten im Zivilrecht gestellt. Das BGB definiert auch Verjährungsfristen, z. B. für Schadenersatz aus unerlaubter Handlung.
- **Zivilprozessordnung (ZPO)**  
Analog zum BGB wird durch die ZPO geregelt, welche Dokumente als Urkunde anerkannt werden müssen, beispielsweise aufgrund einer eigenhändigen Unterschrift oder einer qualifizierten digitalen Signatur.
- **Handelsgesetzbuch (HGB)**  
Hier werden Anforderungen an die Ordnungsmäßigkeit und Revisionsfähigkeit der Geschäftstätigkeit gestellt. Dies umfasst auch bestimmte Aufbewahrungsfristen für Geschäftsdokumente.
- **Grundsätze ordnungsmäßiger Datenverarbeitung (GoDV)**  
Die GoDV sind selbst keine gesetzliche Vorschrift, sondern hergeleitet aus den im HGB definierten Grundsätzen ordnungsmäßiger Buchführung. Sie sind als de facto-Standard für die DV-Revision in Unternehmen zu verstehen.
- **Grundsätze zum Datenzugriff und zur Prüfbarkeit digitaler Unterlagen (GDPdU)**  
Das Bundesministerium der Finanzen hat die in den GoDV vorgesehenen Revisionsanforderungen im Rahmen der GDPdU präzisiert. Dies betrifft hauptsächlich alle steuerlich relevanten digital vorliegenden Dokumente. Hierbei wird u. a. gefordert, dass alle zur Auswertung der Daten notwendigen Informationen wie Dateistruktur, Datenfelder, interne und externe Verknüpfungen in maschinell auswertbarer Form zur Verfügung stehen müssen.
- **Gesetze und Vorschriften zum Schutz personenbezogener Daten**  
Sofern personenbezogene Daten archiviert werden, müssen die hierfür geltenden Gesetze und Vorschriften eingehalten werden. Dazu gehören vor allem das Bundesdatenschutzgesetz (BDSG) und die entsprechenden Gesetze der Länder.

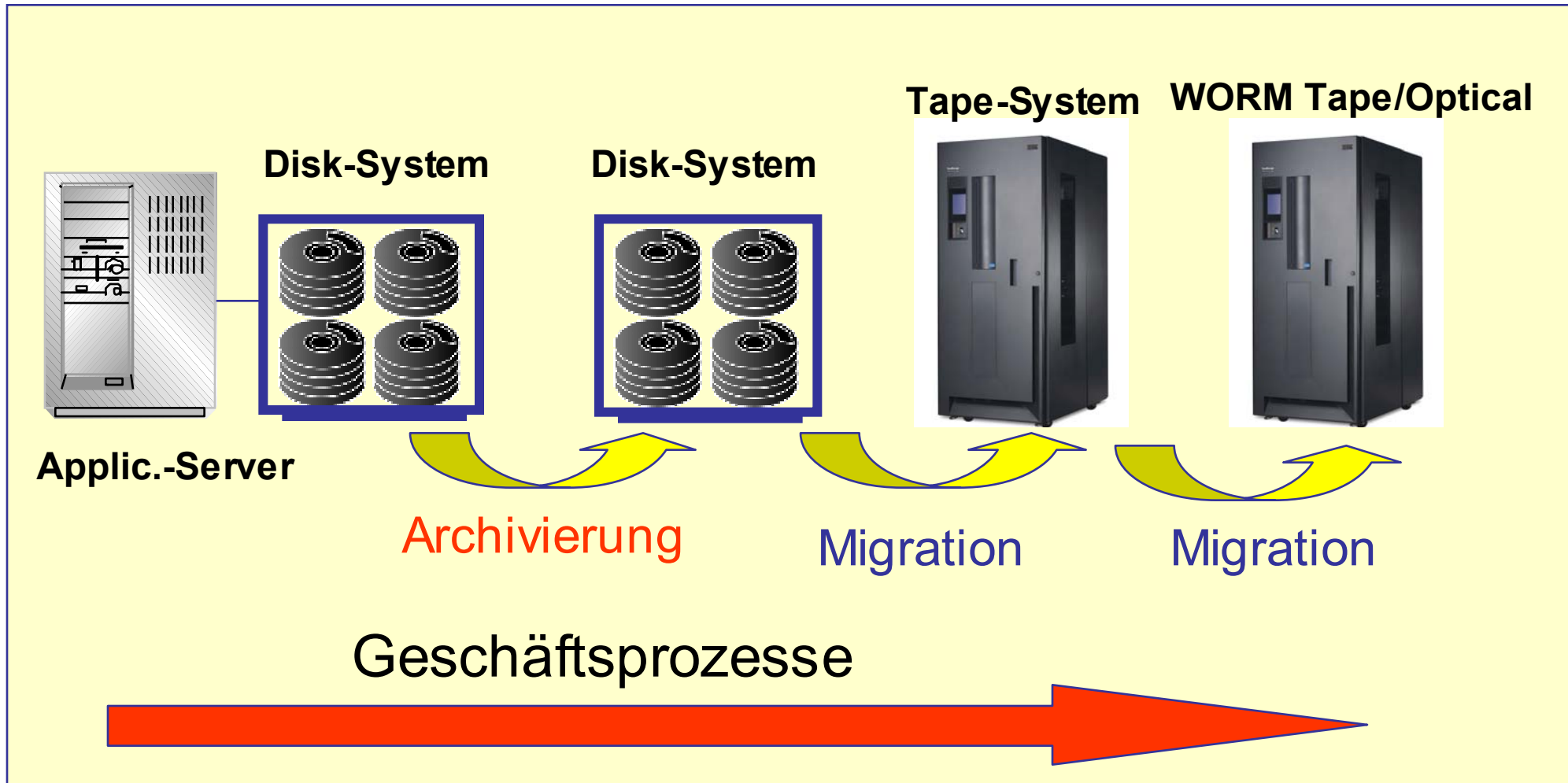
## Gesetzliche Anforderungen

- Gesetzgeber fordert nur, dass durch das Verfahren der Datenspeicherung die Daten
  - ▶ nicht gelöscht oder
  - ▶ verändert werden können
  - ▶ **Nicht aber, auf welches Speichermedium gespeichert werden soll**
    - Solange die Daten les- und auswertbar gespeichert werden
    - ( Seit 1. Januar 2002 der „Grundsätze zum Datenzugriff und zur Prüfbarkeit digitaler Unterlagen (GDPdU)
- Das bedeutet, gesetzliche Anforderungen zur Datenspeicherung erlauben neben
  - ▶ **HW WORM**  
auch
  - ▶ **SOFT WORM (Funktionalität wird mittels Software gewährleistet)**

Es muss also kein optisches Speichersystem sein!

## Gewünschte Archivierungslösung

Flexibel durch „Hierarchisches Speichermanagement“  
und automatisierte Migrationsmöglichkeit



Beispiel

# Information Lifecycle Management

# Information Lifecycle Management

## *Definition und Anforderungen*

- **Was bedeutet Information Lifecycle Management (ILM) ?**
  - **ILM bildet Geschäftsprozesse ab und besteht aus Strategien und Vorgehensweisen für die organisatorische und technische Umsetzung**
  - **ILM umfasst die Verwaltung von Informationen über deren gesamten Lebenszyklus**
  - **ILM ist kein „fertiges Produkt“, sondern ein Gesamtlösungs-Paket**
  - **Ziel ist, Informationen im Unternehmen optimal, effizient und kostengünstig zu nutzen**

# Information Lifecycle Management

## *Definition und Anforderungen*

- **ILM lässt sich in zwei Teilbereiche aufgliedern:**
  - 1. Organisatorisch**  
Einteilung der Daten in Klassen entsprechend ihrer Wichtigkeit zu den Geschäftsprozessen im Unternehmen
  - 2. Technisch**  
Optimale Speicherung der Informationen in die zugewiesene Datenklasse. Eine Datenklasse repräsentiert eine dedizierte Speicherklasse bzw. eine Speicherarchitektur

# Information Lifecycle Management

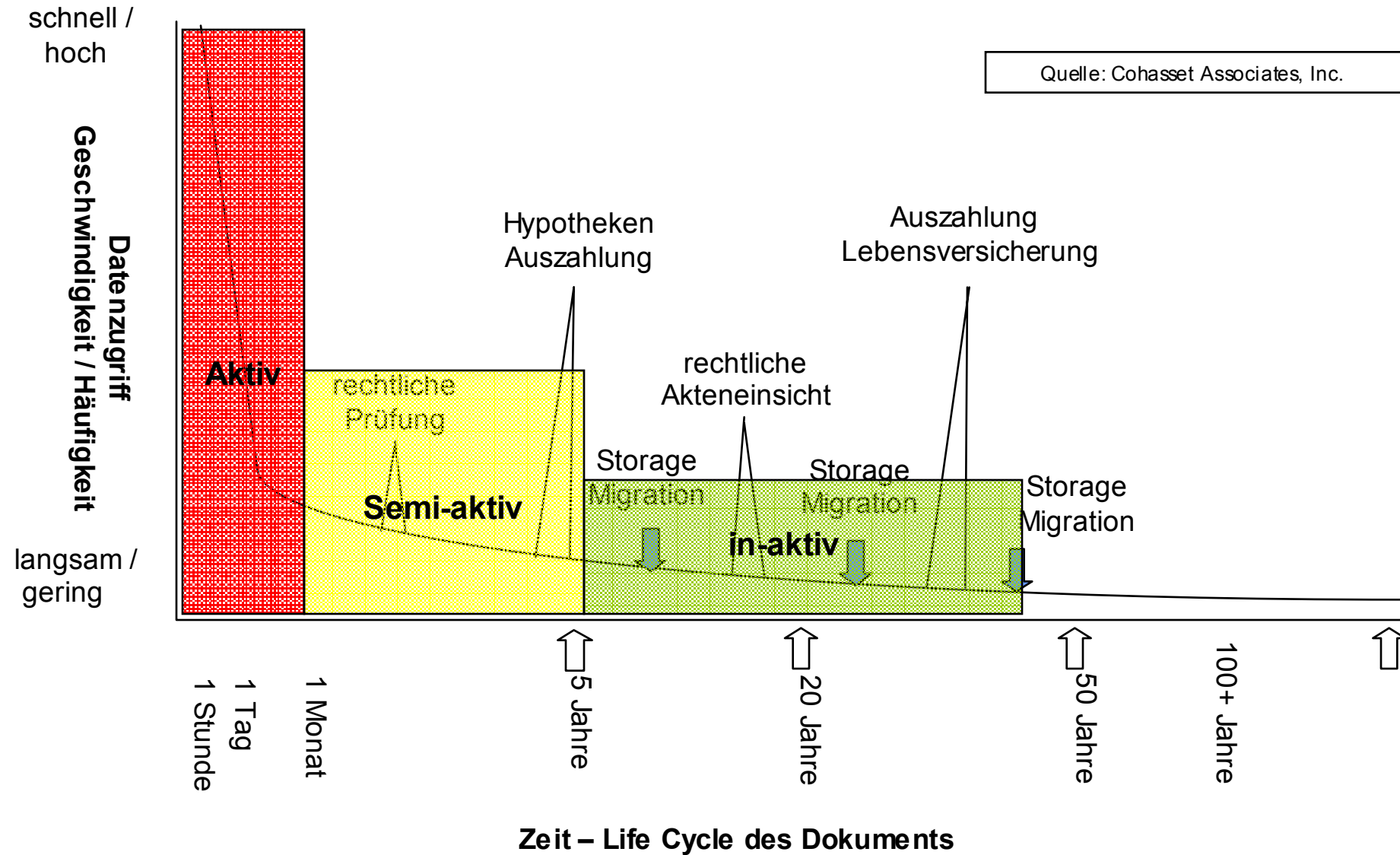
## *Definition und Anforderungen*

- **Wesentliche Komponenten bei der Umsetzung von ILM**
  - **Konzeption, nach der die Klassifizierung und Migration von Daten klar definierten Geschäftsprozessen-und Regeln unterliegt**
  - **Einsatz geeigneter Anwendungslösungen (DMS, Resource-Mgmt, Hierarchisches Speicher- Mgmt. usw)**
  - **Etablieren einer geeigneten Speicherarchitektur- und Hierarchie**
  - **Einbinden von Backup/Restore, Archivierungs- und Disaster-Recovery-Lösungen in ein ILM-Konzept**



# Information Lifecycle Management

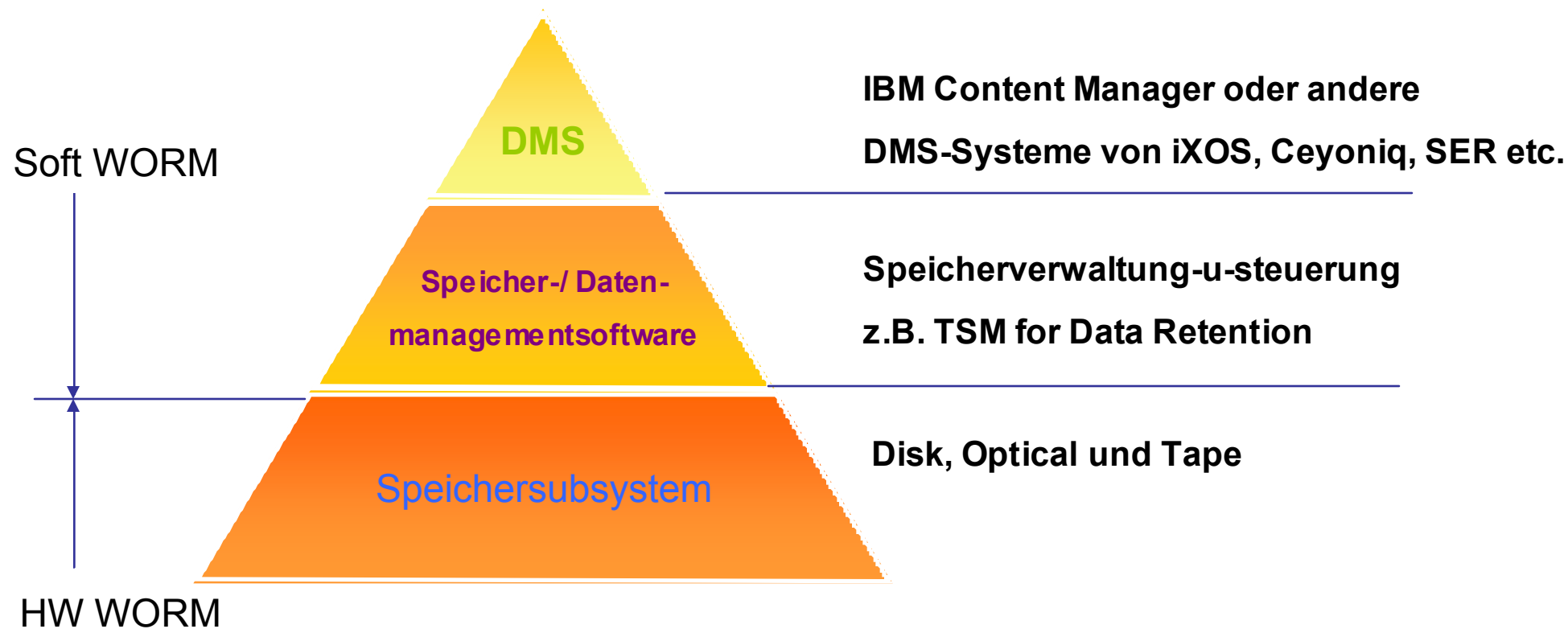
Zugriffsanforderung über den Archivierungszeitraum



Archivierungslösung  
mit  
IBM TotalStorage Data Retention 550

# Archivierung

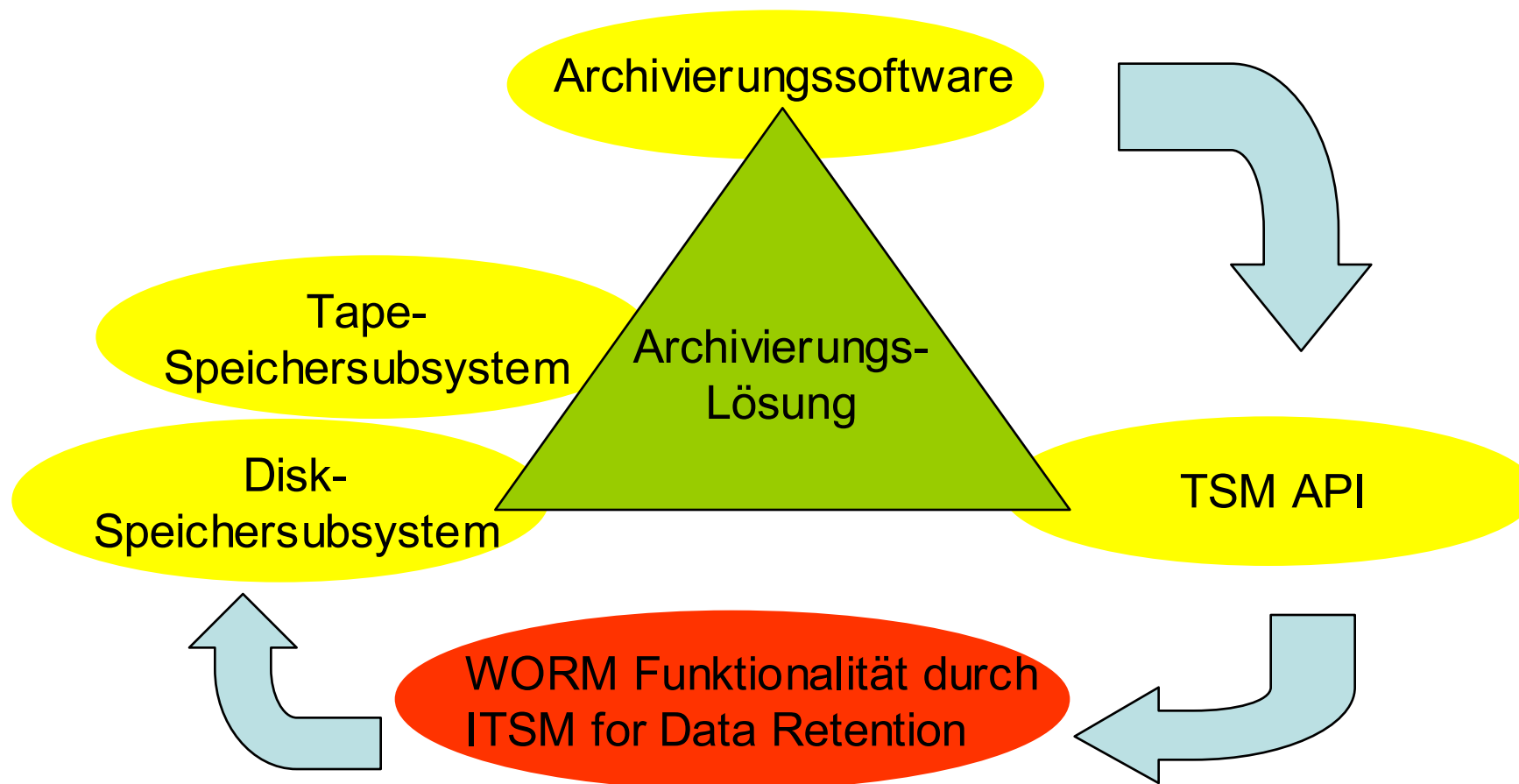
## Komponenten einer Archivierungslösung



- **Unveränderbarkeit der archivierten Daten garantiert durch**
  - ▶ Dokumentenmanagementsystem und/oder
  - ▶ Speicher-/Datenmanagementsoftware und/oder
  - ▶ Speichersystem (Speichersubsystem)

## WORM-Funktionalität auf Disk- und Tapesystemen

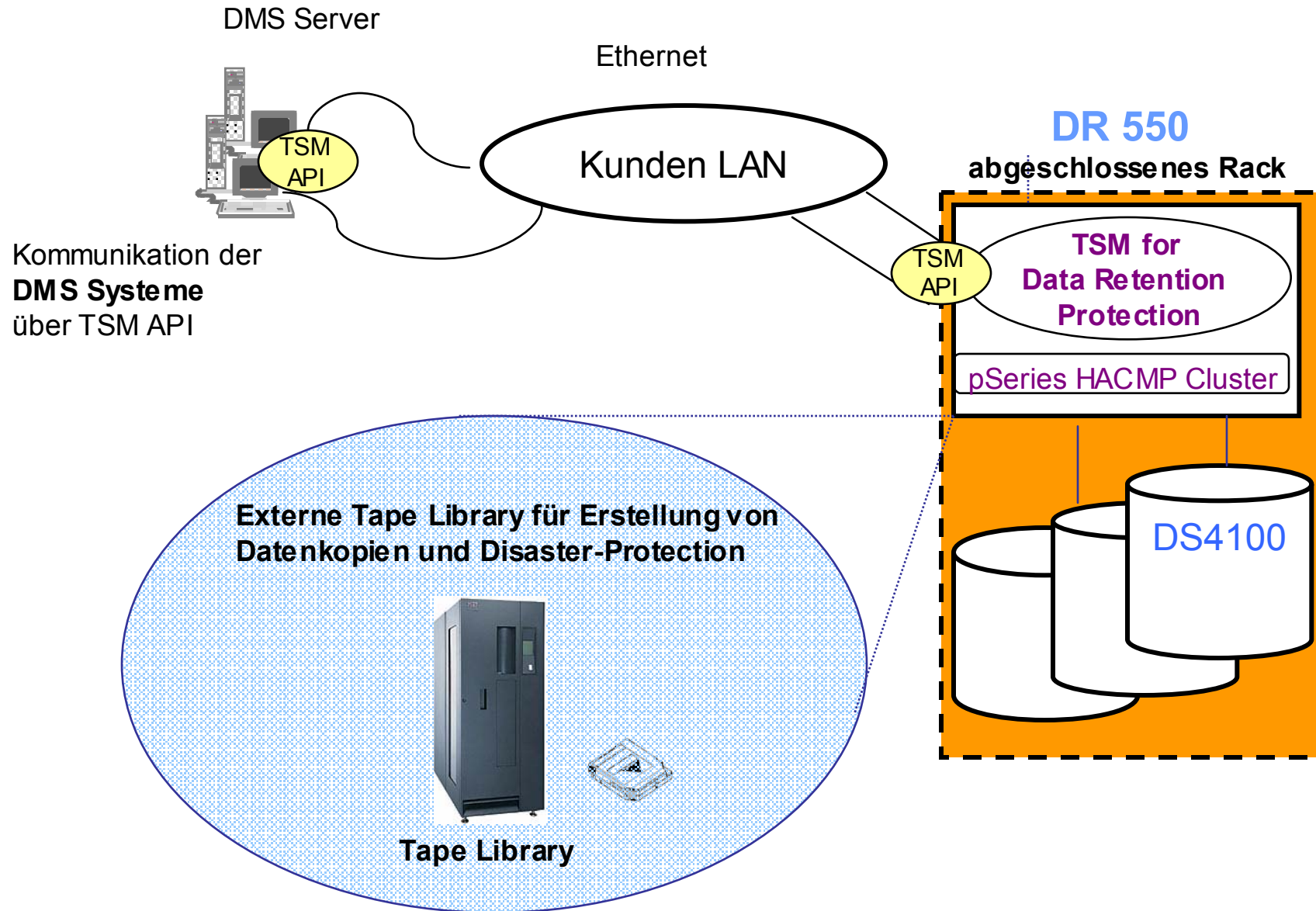
- **IBM Data Retention Solution ermöglicht WORM-Funktionalität auf Plattensystemen und Bandsystemen**
  - ▶ Durch Softwarefunktionalität mittels “IBM Tivoli Storage Manager for Data Retention”



## TSM for Data Retention

- TSM ist zentraler Bestandteil zur Sicherstellung der WORM Funktionalität
- TSM for Data Retention Protection beinhaltet folgende Grund-Funktionen:
  - ▶ Archivierte Objekte können **nicht** vor Ablauf des definierten Archivierungszeitraumes **gelöscht oder anderweitig verändert werden**
  - ▶ Festlegen der Verwaltungsregeln über Retention/Delete Policy.
  - ▶ „Retention Policy“: Retention Periode kann gestartet werden
    - Zum Zeitpunkt der Archivierung (Chronologische Archivierung)
    - Aufgrund von Ereignissen (Event-basierende Retention)
  - ▶ „Delete Policy“:
    - Eintritt des Ablaufdatums: Daten werden automatisch gelöscht (Data - shredding)
    - Deletion Hold: Retention Daten werden, obwohl diese mit Ablaufdatum gekennzeichnet sind, gehalten und nicht gelöscht.
    - Deletion Release: Retention Daten werden wieder freigegeben (Zum Löschen)

# Architekturübersicht der DR550



## IBM TotalStorage Data Retention 550

- **Herzstück der Lösung ist *TSM for Data Retention***
  - ▶ archivierte Dokumente können nicht modifiziert bzw. gelöscht werden
    - Daten werden auf Platte (Raid-5 od. Raid-10) gesichert
  - ▶ Integriertes hierarchisches Speichermanagement
    - Kopie oder Migration auf Tapes
  - ▶ Integriertes Retention Management
- **Plattenkapazität bis 56 TB (DS4100)**
- **Hochverfügbarkeit durch:**
  - ▶ Zwei pSeries 520 als HACMP Cluster
  - ▶ Redundanter Netzwerk- und Stromanschluss
- **Flexibel durch Single-Node Konfiguration**
- **Anschluss an Kundennetzwerk über TCP/IP (Ethernet)**
  - ▶ Kommunikation mit DMS-Systemen über TSM API
- **Erweiterung mit externer Devices möglich**
  - ▶ 3592 Tape-Drives (WORM Schutz) oder Disk
  - ▶ Andere von TSM unterstützte Devices

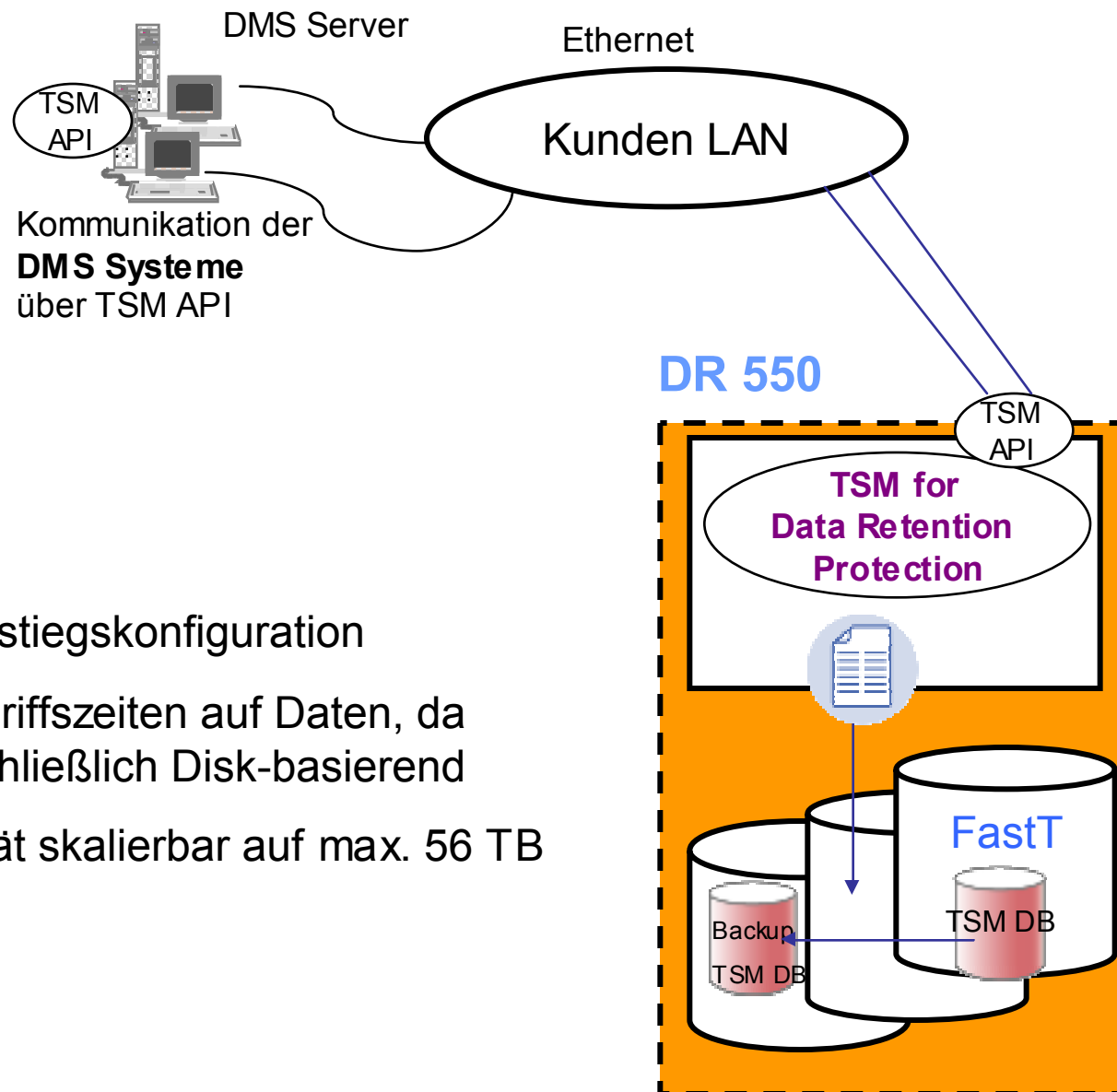


# Lösungsbeispiele



# IBM TotalStorage Data Retention 550

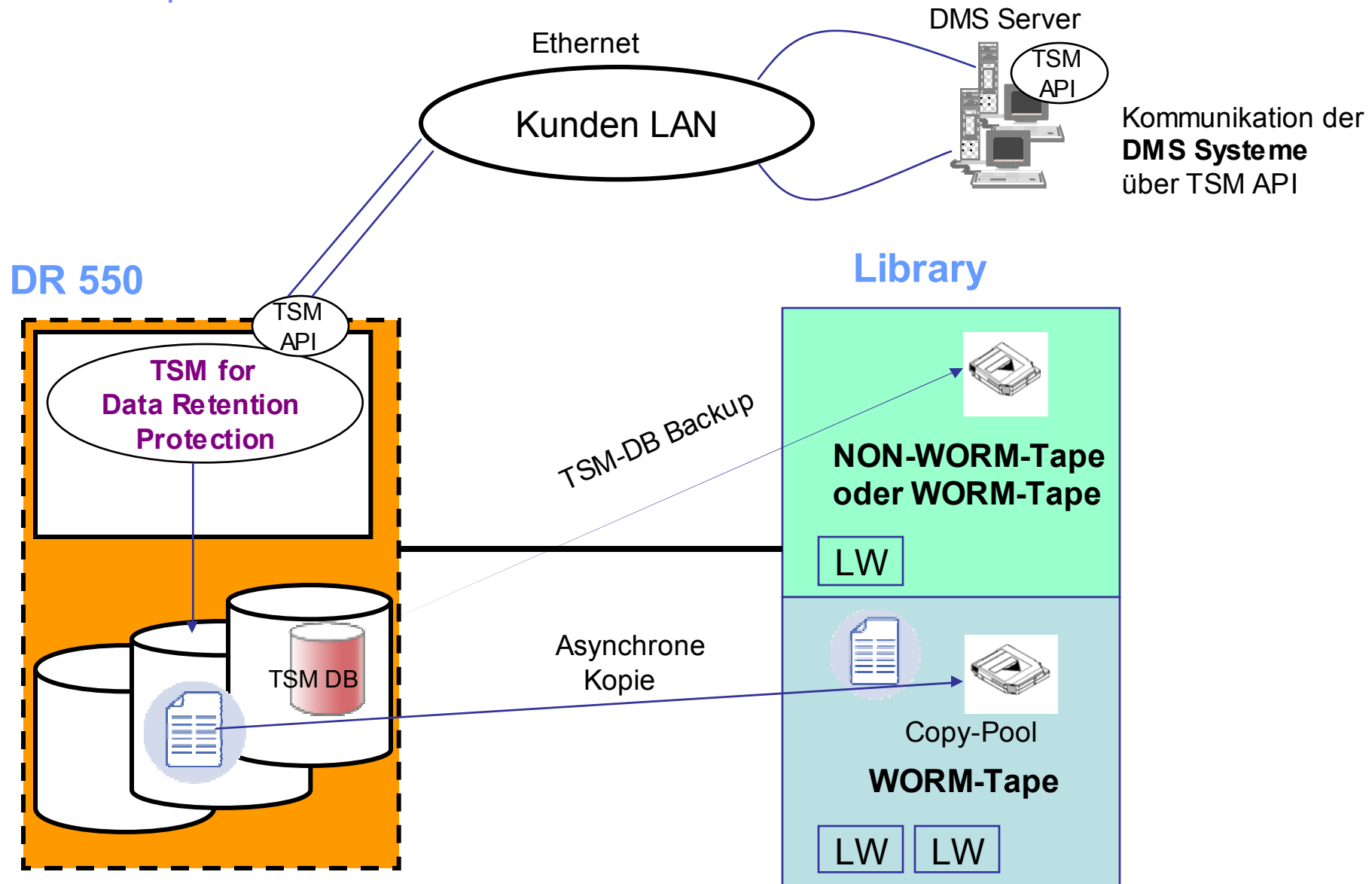
Lösung 1 (Basis): Archivieren auf interne Platte einer DR550.



- Günstige Einstiegskonfiguration
- Schnelle Zugriffszeiten auf Daten, da Lösung ausschließlich Disk-basierend
- Disk-Kapazität skalierbar auf max. 56 TB

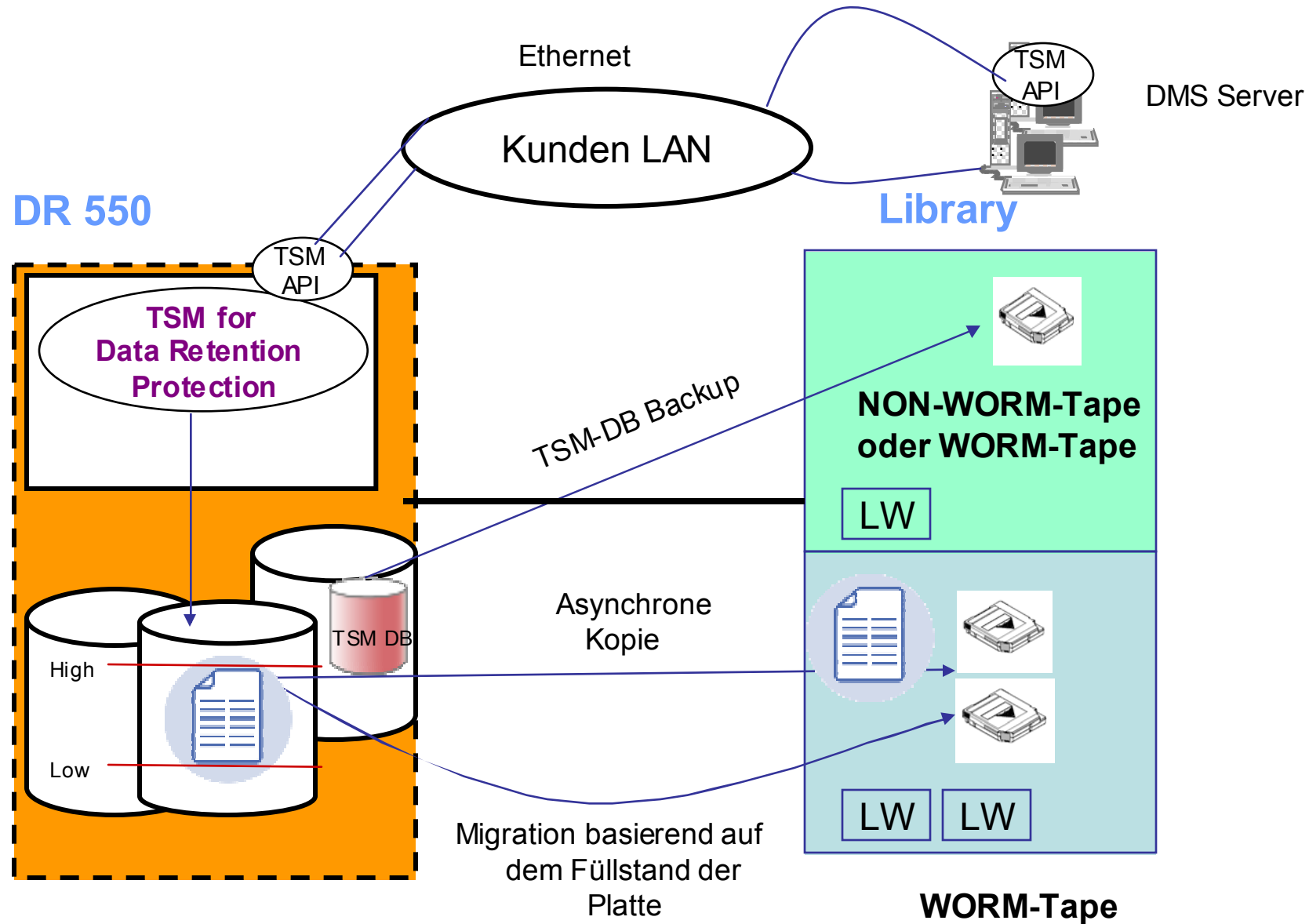
# IBM TotalStorage Data Retention 550

Lösung 2a: Archivieren auf interne Platten der DR 550. Asynchrones kopieren auf WORM Tape



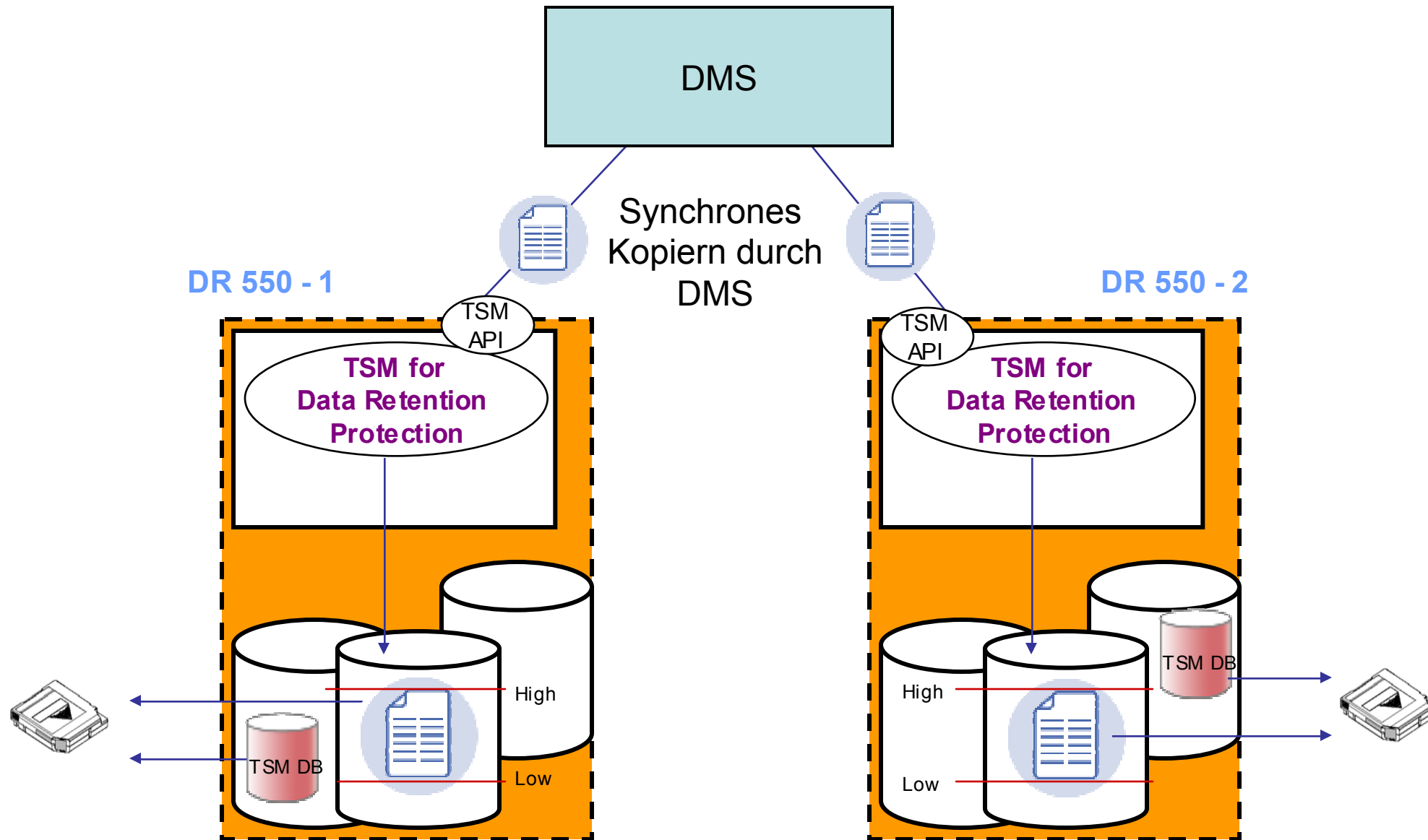
# IBM TotalStorage Data Retention 550

Lösung 2b: Archivieren auf interne Platten, Migration und Backup auf WORM Tape, Backup DB auf Non-WORM-Tape oder WORM-Tape)



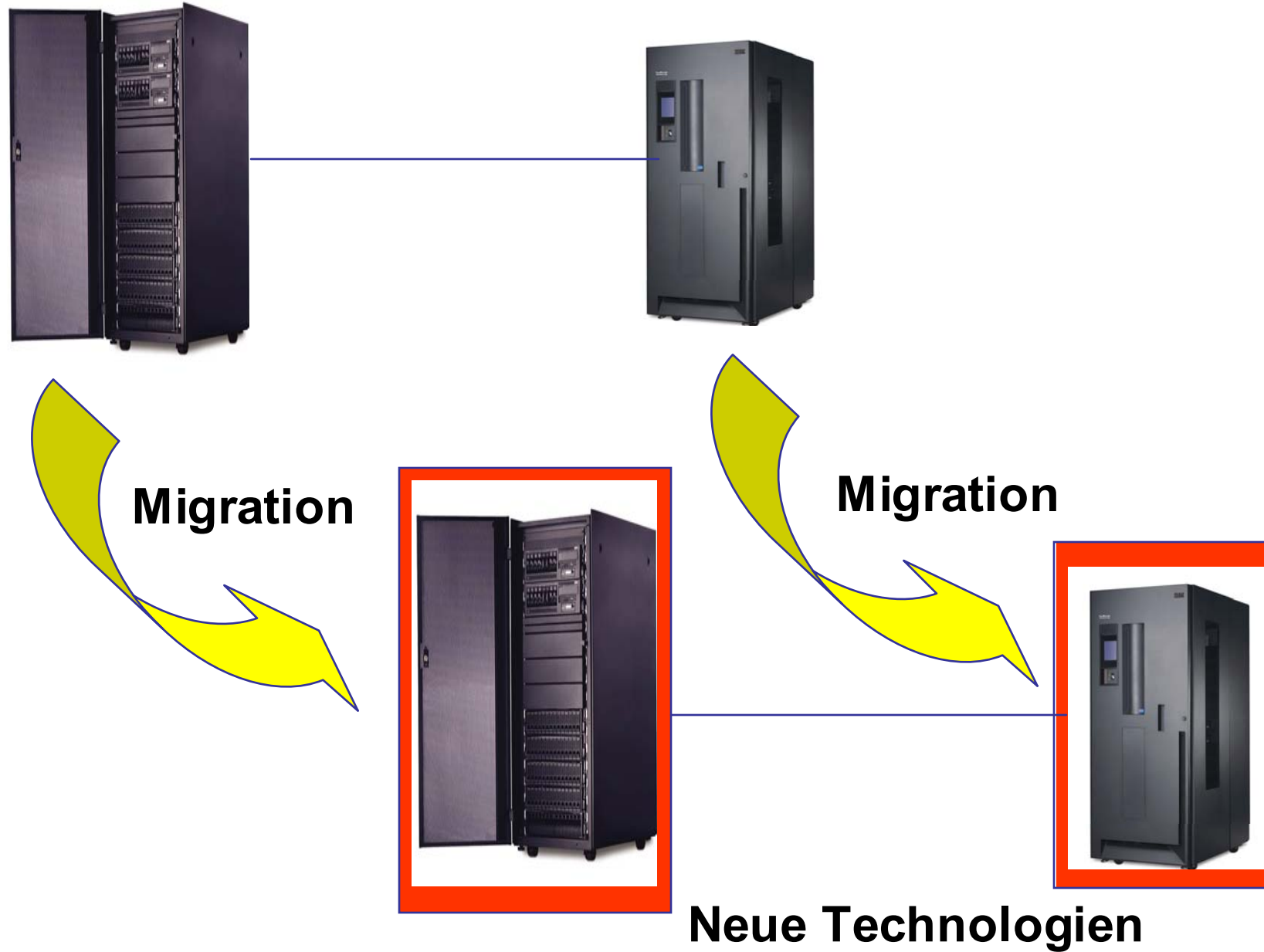
# IBM TotalStorage Data Retention 550

Lösung 3: Synchrones Kopieren mittels Dokumenten Management System - Funktionalität



# IBM TotalStorage Data Retention 550

## Einfache Migration in zukünftige Technologien



## Das Jahr des Jaguar: 3592 Generation 2



- **Neue Flat Lap Kopf Technologie Generation 2**
- **500 GB native Kassettenkapazität**
- **16-Spur Technik**
- **100 MB/S native Datenrate**
- **896 Spuren**
- **schnellste Spulgeschwindigkeit mit 10 m/Sek.**
- **Noch höhere Datensicherheit durch weiterentwickelten Error Correction Code**
- **Schnellerer Load-/Unload-Mechanismus**
- **Unterstützt Gen.1 und Gen.2 Modus**
- **Extended Length-Kassetten in 2006 mit +700GB native**

**Schnellstes Laufwerk weltweit**  
**Erstes Laufwerk mit 4 Gbit FCP**

HighTech auf kleinstem Raum

# Zusammenfassung

## ***IBM Archivierungslösung mit DR550***

### ***Vorteile***

- **Flexible Lösung durch integriertes “Hierarchisches Speicher Management (HSM)” nach Wert der Information**
- **Kostengünstige und kundenindividuelle Lösungen**
- **Migration von Archiv-Daten auf zukünftige Speicher-Technologien durch TSM-Migrations-Funktionalität problemlos realisierbar**
- **Zertifizierte Lösung**
- **Hochverfügbar**
- **Unterstützt alle Anforderungen hinsichtlich:**
  - ▶ **Zuordnung von Speicher analog der Geschäftsprozesse (ILM)**
  - ▶ **Rechtliche Anforderungen**
  - ▶ **Aufbewahrungsdauer und Zugriffshäufigkeit- und Geschwindigkeit**