

Oxford University TSM Symposium 2005

Tivoli Storage Manager : Facing the Future  
St Catherine's College, 27-29 September 2005



IBM TotalStorage Software

# Aspects of Information Lifecycle Management

TSM and Retention Managed Data



Klemens Poschke  
IT Architect IBM Softwaregroup  
klemens.poschke@de.ibm.com



© 2005 IBM Corporation

# Task-Definitions

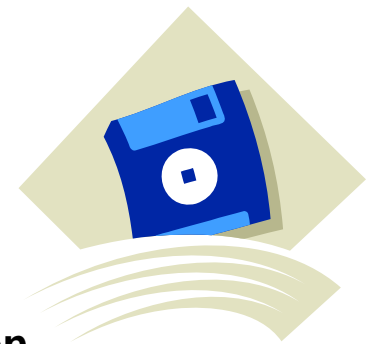
Archive	Backup & Recovery	HSM
compliant original data	restorable copy	Original data
Read-access rare (exception: Data-Mining)	Read-access rare (only Restore)	Many reads
Never overwrite	few overwrites (versioning)	Many overwrites

The task defines the necessary storage resources and the required compliance concept



## Requirements on compliant archiving

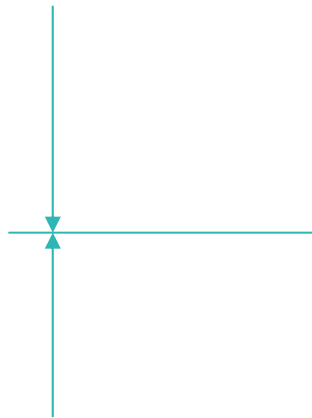
- **Archiving solution must guarantee that data**
  - can not be deleted or modified outside of application control
  - must be kept during defined periods
  - can be read after long times
  - can be protected for disaster recovery
  
- **Storage media is not specified by compliance regulations**
  - Data must be readable and access must be controllable
  - Ambivalence of long-time retention and technological progress
  
- **Evaluation of long-term cost should be criteria for technology decision**
  - Magnetic disk technology will be substituted within 3 – 5 years
  - Magnetic tape can store data more than 20 years, however downward-compatibility is less (typically 6 years - see example LTO)
  - Optical media currently does not have a trustable technology perspective
  - Cost of technology migration should be reflected in today's investment decisions



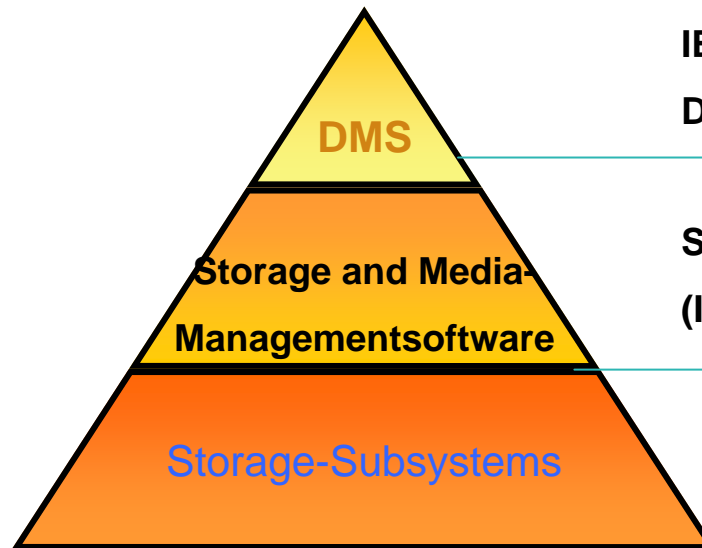
# Archiving

## Components of IBM's archiving solutions

Software WORM



Hardware WORM



**IBM Content Manager or other vendors**

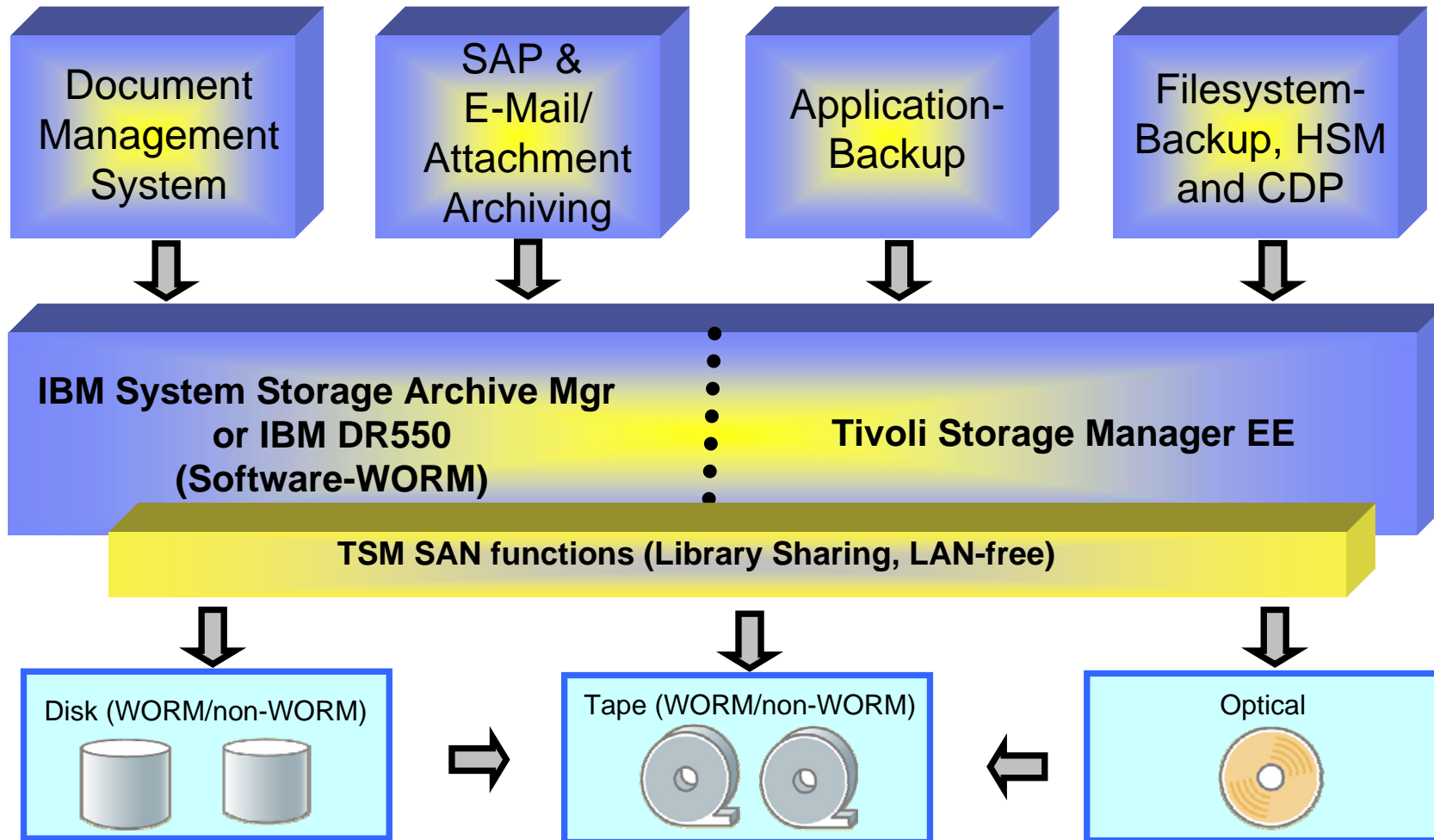
**DMS-Applications (e.g. iXOS, KVS, Filenet)**

**Storage- and Data-Management Application  
(IBM System Storage Archive Manager<sup>\*\*\*</sup>)**

**Disk, Optical and Tape**

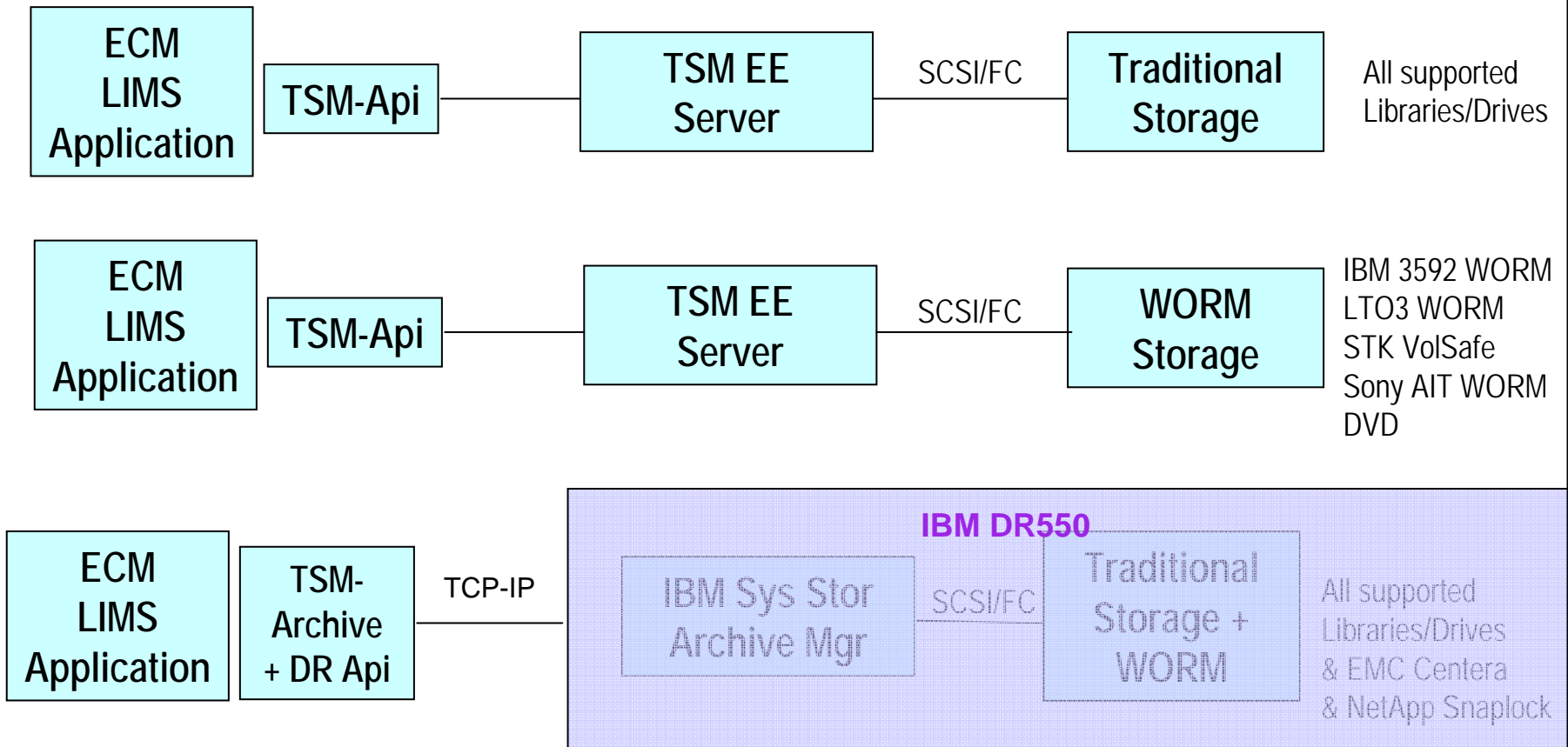
<sup>\*\*\*</sup> name changed from „TSM for Data Retention“ to „IBM System Storage Archive Manager“ on Sep 06,2005

# TSM is the universal Media Manager for Backup and Archiving





# Data Retention Solutions with TSM



## Characteristics and functions of „Archive Manager“

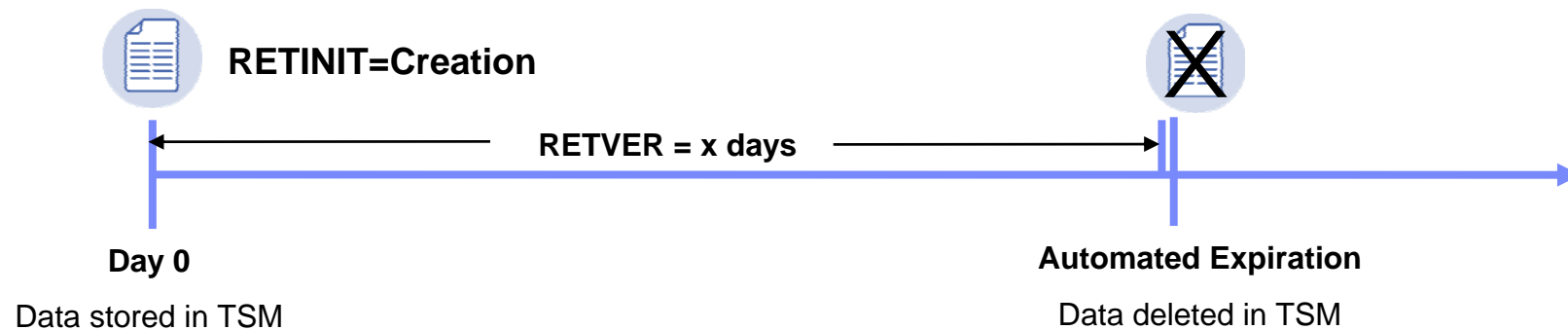
- **„Archive Manager“ is only usable for archiving**
  - dedicated „Archive Manager“ Server
  - accessible by TSM Archive API ( $\geq$ V5.2.2) or TSM Archive Client ( $\geq$ V5.3.2)
  - does not allow for TSM Backup API, HSM or Backup-Client communication
  
- **„Archive Manager“ new functions**
  - different „Retention Policies“ can be defined on the Server:
    - Chronological – Retention starts when writing archive object in TSM
    - Event-based – Retention time is indefinite and ends with initiated „event“ from client
  - „Delete Policy“ initiated on the API Client (by the application):
    - Send „event“-call to archive object to signal end of retention („delete object“ is not supported)
    - Deletion Hold: data will be retained until „deletion release“
    - Deletion Release: give back control to normal retention procedures
  - by end of retention period or initiated „event“ data will be deleted in TSM by normal expiration process





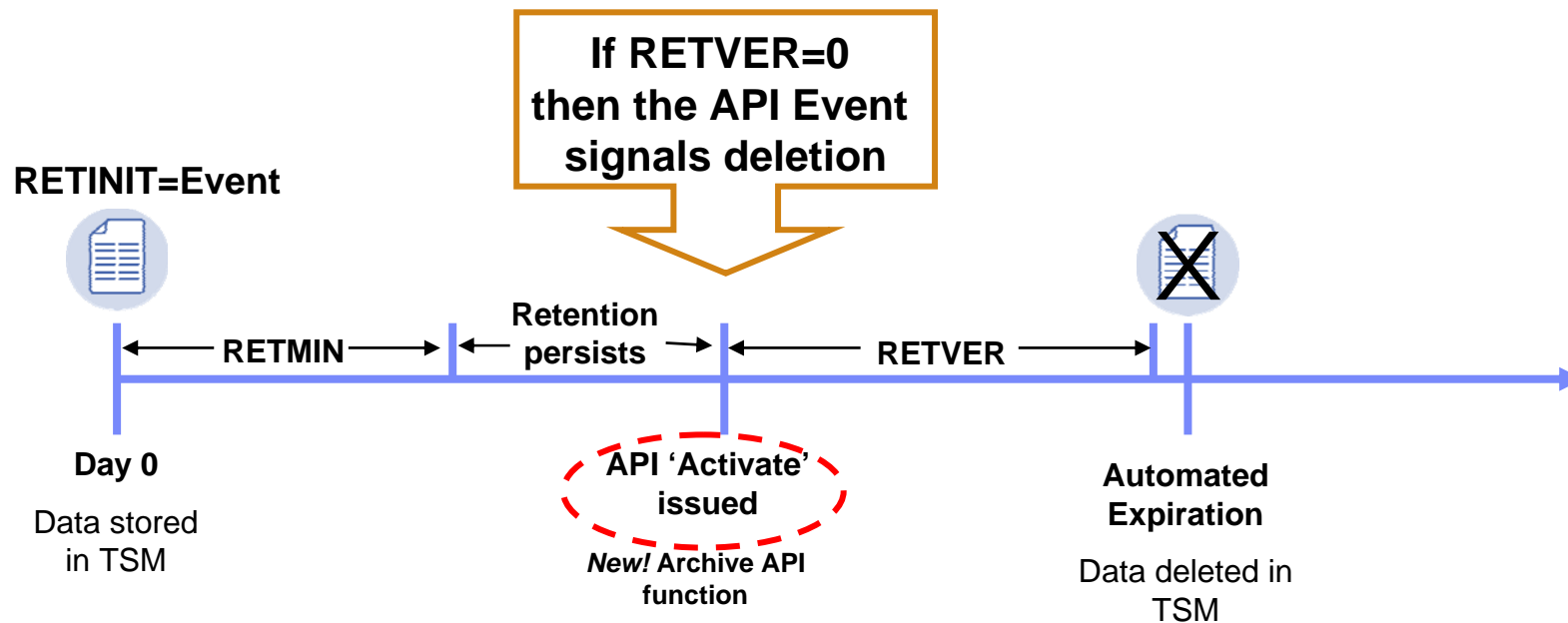
## Chronological Retention Policy

- Data retention starts when object is stored on TSM Server
- Setup Chronological retention using parameter `RETINIT=CREATION`  
`RETVR` specifies the time to retain objects after it is stored



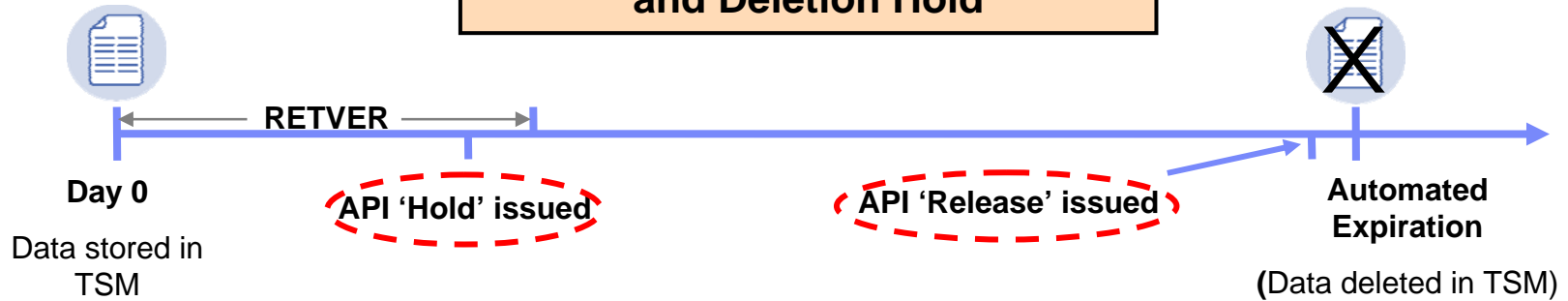
## Event Based Retention Policy

- **Event from API signals start of retention time or deletion**
  - Start counting down retention clock at occurrence of event
  - Only one event for each object
- **Setup Event-based retention using parameter RETINIT=EVENT**
  - RETMIN specifies the minimum time to retain object after creation
  - RETVER specifies the time to retain objects after event occurred
  - Expiration always on later criterion (RETMIN or event + RETVER)

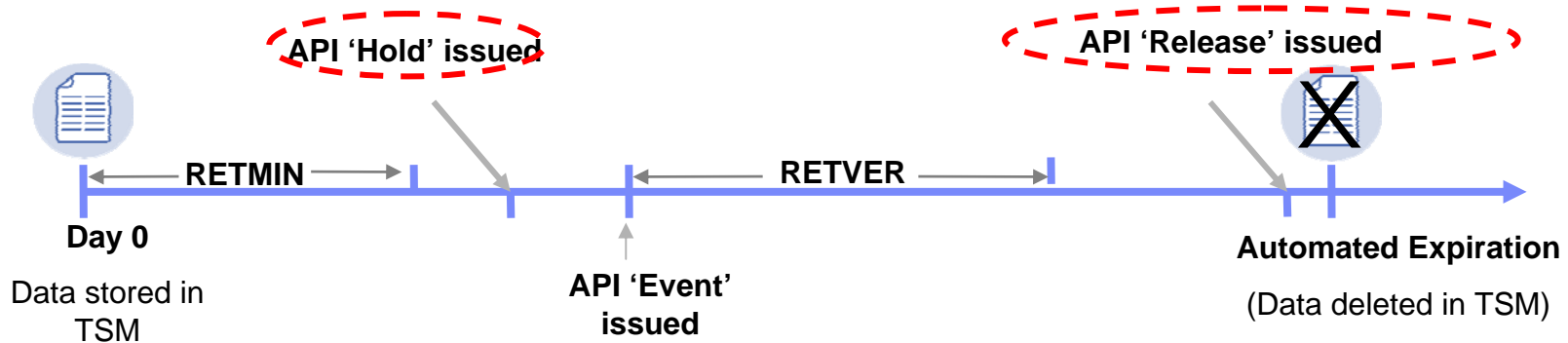


# Data Retention Policy Summary

## Chronological Retention Policy and Deletion Hold

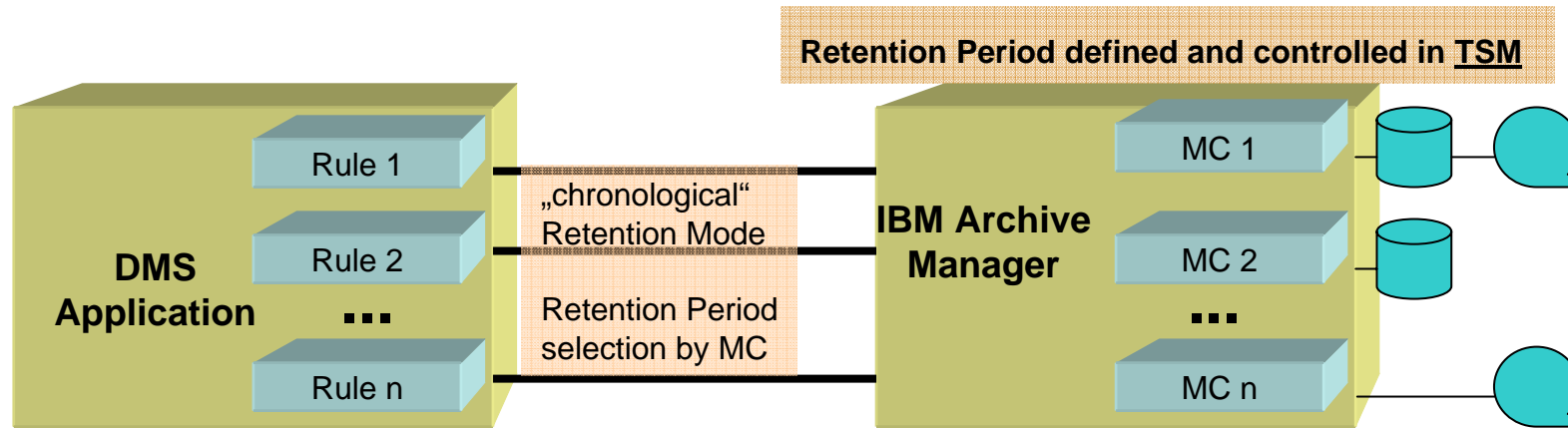


## Event-Based Retention Policy and Deletion Hold

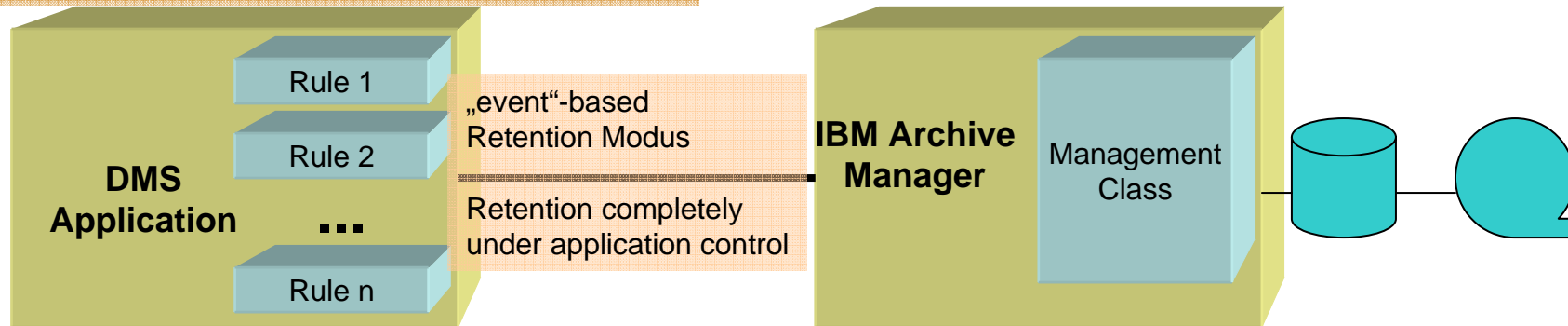


# Retention-Methodology

## Server-oriented vs. Application-oriented



**Retention Period defined and controlled in Application**



## Examples from a Retention Architecture

**Retention Period INDEFINITE = document can be deleted at any given time by a client request**

Implementation: RETINIT=EVENT, RETMIN=0, RETVER=0. If the client application intends to delete the document it starts the ACTIVATE\_EVENT and the document is deleted immediately

**Retention Period FOREVER**

Implementation: RETINIT=CREATION, RETVER=NOLIMIT

**Retention Period FIXPERIOD Begin IMMEDIATE = document is stored for a fixed amount of time, retention clock starts immediately**

Implementation: RETINIT=CREATION, RETVER=<Period> - during document generation an immediate HOLD-Event will be raised - the RELEASE will define the deletion point if RETVER has been exceeded

**Retention Period: FIXPERIOD Begin EVENTBASED = document will be retained a fixed number of days, retention clock starts when application raises EVENT**

Implementation: RETINIT=EVENT, RETMIN=<Period>, RETVER=<Period> - document is stored forever until EVENT is activated, at the same time a HOLD is raised



# IBM Archive Manager

## *Ensure Software-WORM*

- **TSM „archive" function does not allow overwrite at all, many identical archive objects can be stored in parallel**
  
- **Archive Manager restricts admin operations**
  - delete an archive object (before RETVER has been exceeded) – **not allowed**
  - reduce retention periods - **not allowed**
  - **DELETE FILESPACE - not allowed**
  - **DELETE VOLUME DISCARDATA=YES - not allowed**
  - **AUDIT VOLUME FIX=YES - not allowed**
  - Reassign node to a different policy domain – **not allowed**
  - device class with a device type of **SERVER** – **not allowed**
  - import data to a retention protected server – **not allowed**
  - You can export data but it will not be retention protected once it is imported

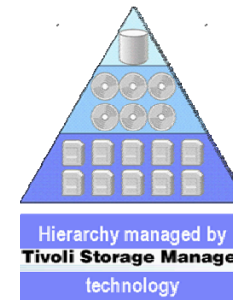
These restrictions guarantee that TSM-Admin can not delete archive data on physical resources or modify retention associations to stored data.



# IBM Archive Manager

## Packaging

- **For any supported operating system and storage configuration**
  - IBM System Storage Archive Manager Software
  - All TSM Server-Platforms
  - all supported devices on those platforms



- **Packaged solution: IBM DR550**
  - TotalStorage Data Retention 550
  - pSeries Power5 with/without Cluster-Support, DS4100 S-ATA in a lockable rack
  - Restrictions for AIX and DS4000 Admin on physical configuration
  - can be easily expanded with tape technology
  - selectable redundancy configurations
  - certified by public accountant (in Germany and Switzerland)



**Single Node or Cluster**

- **3,5 TB (1 EXP. 100)**
  - ▶ 1 EXP 100 (14 SATA Disk)
- **7 TB (2 EXP. 100)**
- **14 TB (4 EXP. 100)**
- **24,5 TB (7 EXP. 100)**
- **56 TB (16 EXP. 100)**

**NEW: DR550 Express Pak**

- p520 with 1TB Disk
- Rack mountable



IBM TotalStorage Software

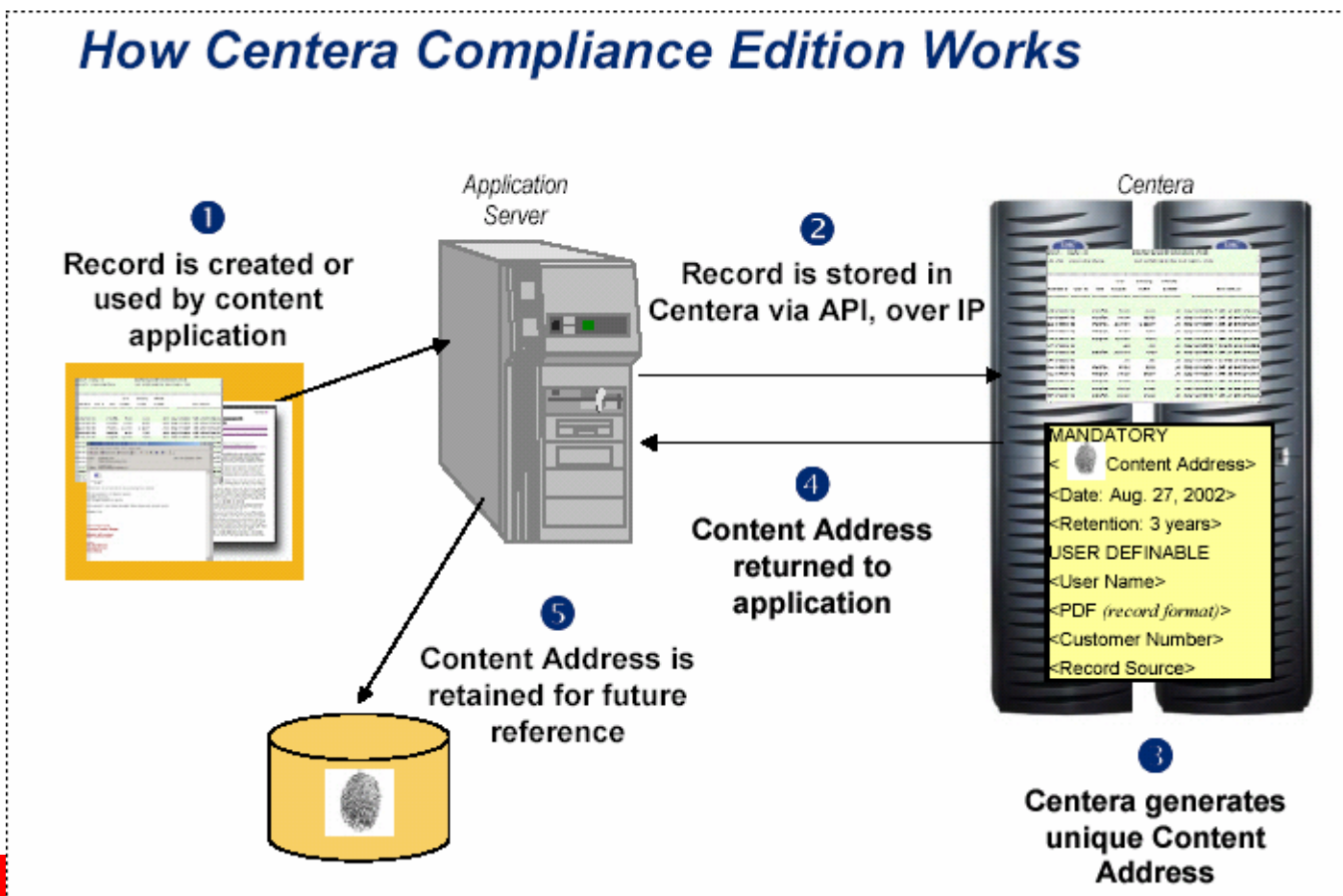
# TSM Operation with EMC Centera and NetApp Snaplock





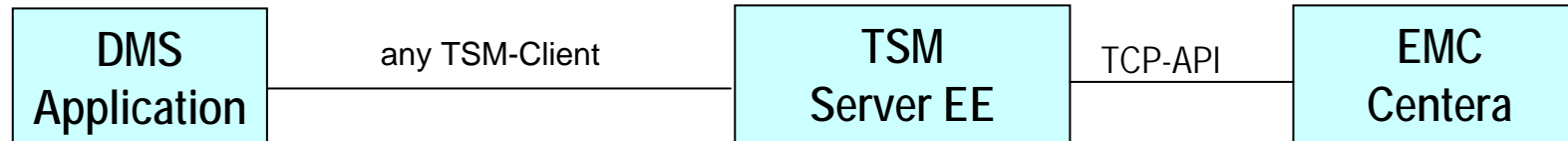
# EMC Centera

- CAS = Content Addressable Storage  
 Content address is a unique fingerprint identifying the data  
 Content address serves as a reference to data



Source: "A Centera Whitepaper: Understanding and Addressing the Challenges of Compliance and Discovery" by EMC Corp.

## TSM EE vs Archive Manager and Centera



- With TSM Extended Edition there is no retention protection, i.e. retention value sent to Centera is 0
- data is not protected against TSM admin delete operations



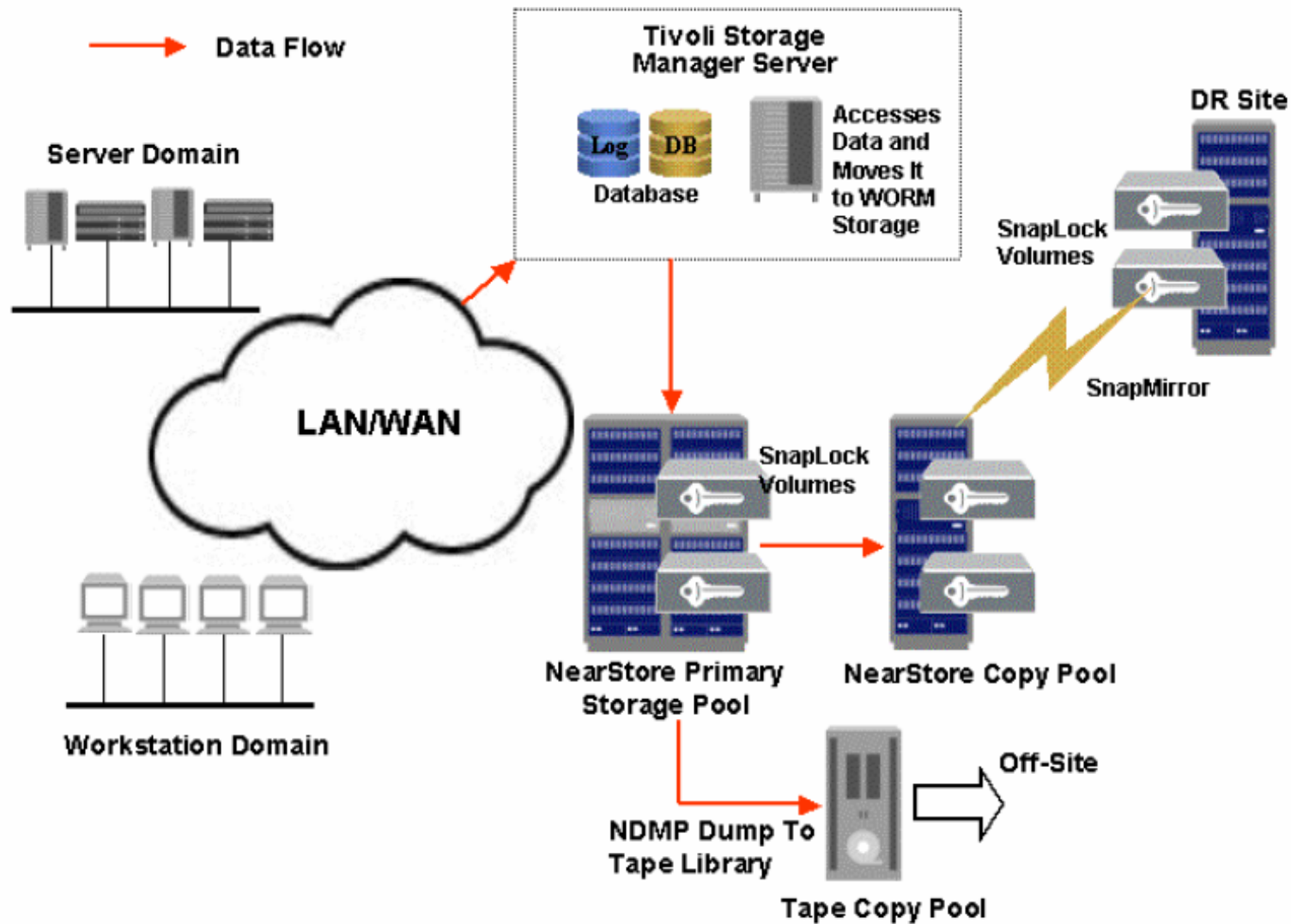
- With Archive Manager there is retention protection, (Archive Mgr sends the retention time to Centera RETINIT=EVENT, RETMIN=retention time, RETVER=0)
- Centera does not automatically expire data, but allows deletion after expiration
- Archive Manager protects against admin deletion of physical resources
- Upon expiration Archive Manager will expire the indices and delete the data on Centera
- Space on Centera can be re-used

## TSM & Centera: Unsupported Operations

- „migration“
- „reclamation“
- „move data“ into or out of a Centera storage pool.
- „move nodedata“ into or out of a Centera storage pool.
- „backup storagepool“ of Centera storage pools.
- „restore volume“ of Centera storage pool volumes.
- **Exporting** data to a Centera device class or **importing** data from a Centera device class; however, files stored in Centera storage pools can be exported and files being imported can be stored on Centera.
- Using a Centera device class for creating **backup sets**; however, files stored in Centera storage pools can be sent to backup sets.
- Defining Centera volumes.
- Using a Centera device class to **back up a database**.
- Using a Centera device class for **database loading or unloading**.
- Using a Centera device class as the target of volume history, device configuration, trace logs, error logs, or query output files.



# TSM and NetApp Snaplock

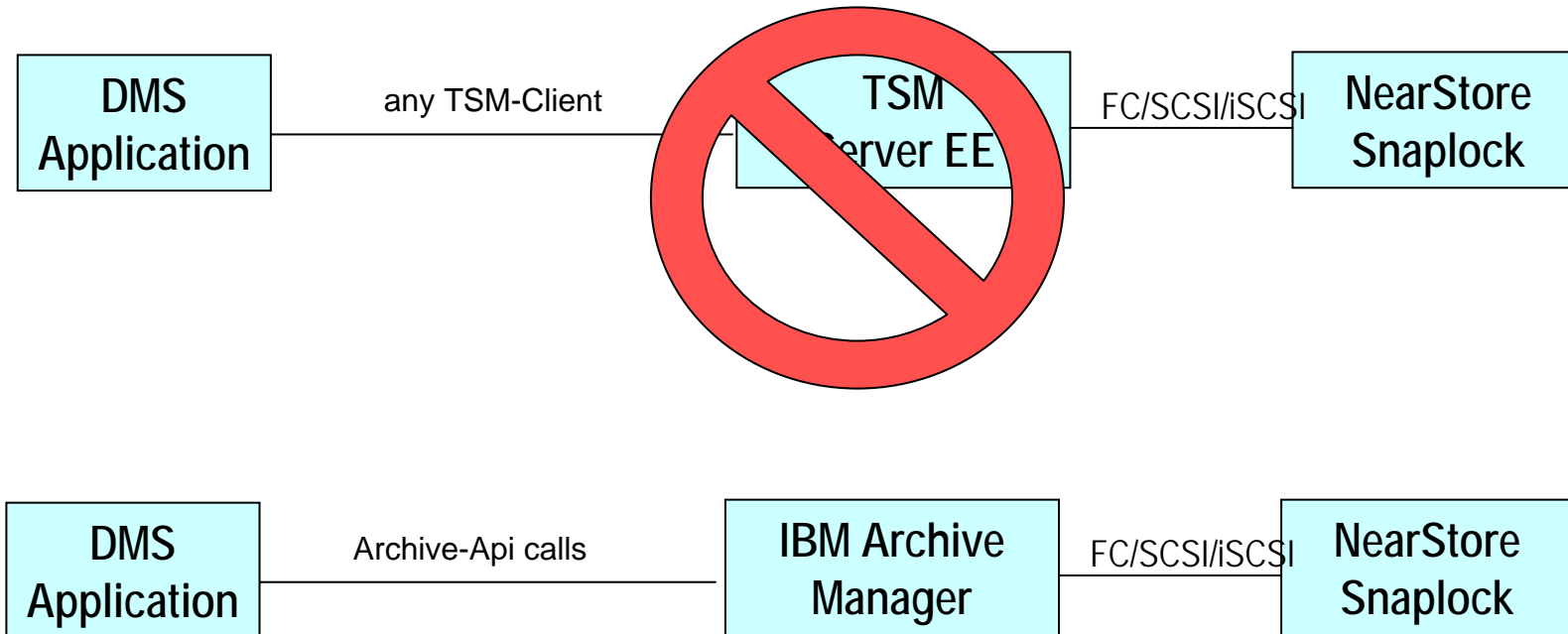


## TSM and NetApp NearStore/Snaplock

- The NearStore/Snaplock system (R200) is managed by TSM as a magnetic sequential file device („file“ diskpool)
- NetApp SnapLock offers write once, read many (WORM) protection for disk-storage
- The data in the primary storage can be dumped to a tape copy pool utilizing NDMP for off-site data protection.
- Secondary storage, or copy pools, can reside on NearStore for data recovery from media in the event of a disaster.
- Alternatively, the copy pool-enabled SnapLock can be subsequently copied using SnapMirror® or SnapVault® to a DR site
- Prerequisites:
  - IBM Archive Manager (not TSM EE!!!) >=V5.3
  - Data ONTAP >=7.1 release



# IBM Archive Server and NetApp Snaplock



- is only supported on Archive Manager Servers (DR550)
- data is stored in WORM File Volumes (seq. file Stgp)
- copy group retention period is to adjust Snaplock WORM File retention period during write transaction
- new Storagepool Parameter RECLAMATIONTYPE=SNAPLOCK
- means reclamation is initiated by a date - not by a threshold
- only empty WORM File Volumes are deleted and space can be re-used

## IBM Archive Manager (DR550) - Highlights

- Synergy with existing or planned TSM Backup environments
  - Skill
  - Infrastructure and storage resources
  - Common Disaster Recovery strategies
- Independance of technological changes for ECM applications
- No need for ECM application to provide device interfaces
- Block-level data-migration between technologies transparent to the application
- Independance from storage hardware supplier
- Expand application functionality by disaster protection capabilities









**And you think you're overworked...**

