

## Information Lifecycle Management



effiziente und kostengünstige  
Datenspeicherung und -archivierung



Distinguished Blue Diamond Award



**FRITZ & MACZIOL : INFOMA®**

# Today's Storage Management Challenges

- exponential growth of storage capacity
- Exponential growth of objects (i.e. datasets) to be managed
- Shrinking backup (and recovery) windows
- Budget cuts
- Regulatory compliance

**ILM addresses all these items**

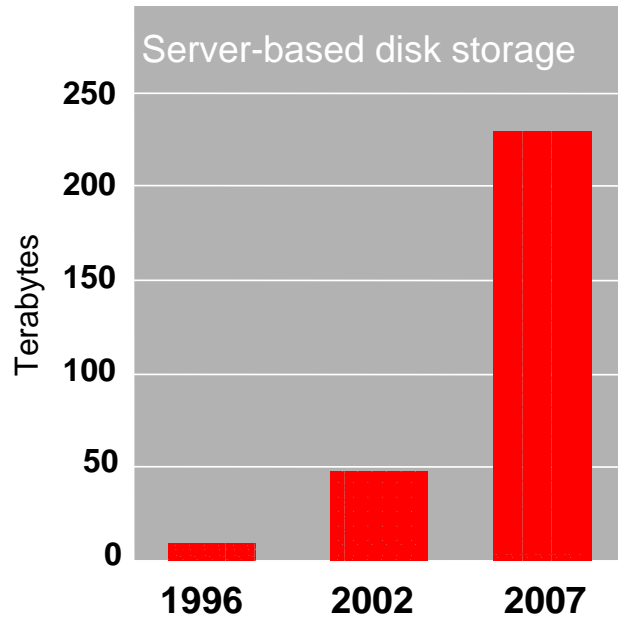
# Information Lifecycle Management

- **ILM is the process of managing information—from creation to disposal—in a manner that aligns costs with the changing value of information**
- **ILM includes Tiered Storage Management for efficiency of storage operations**
- **Data Lifecycle Management (DLM) is a part of ILM**
- **ILM is a combination of DLM and Enterprise Content Management (ECM)**
- **ILM addresses legal requirements (“compliance”)**

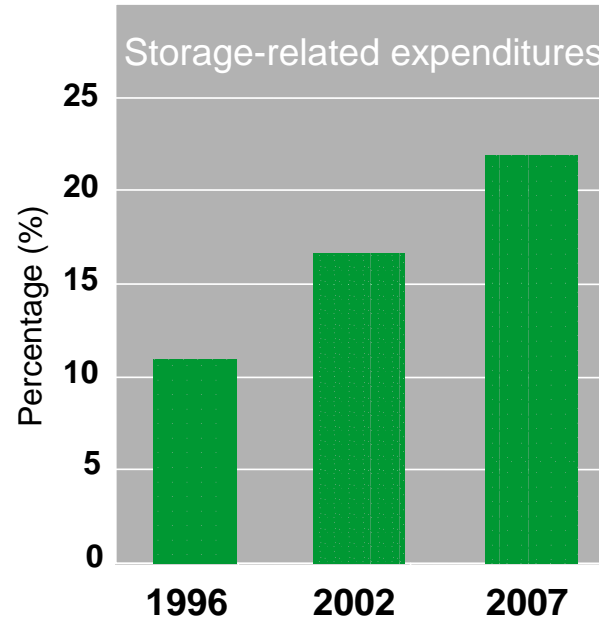
# Storage Management Challenges - Real and Growing

In a typical corporation...

Disk storage is growing rapidly



Storage-related expenditures\*, as a % of IT budgets, is also growing rapidly



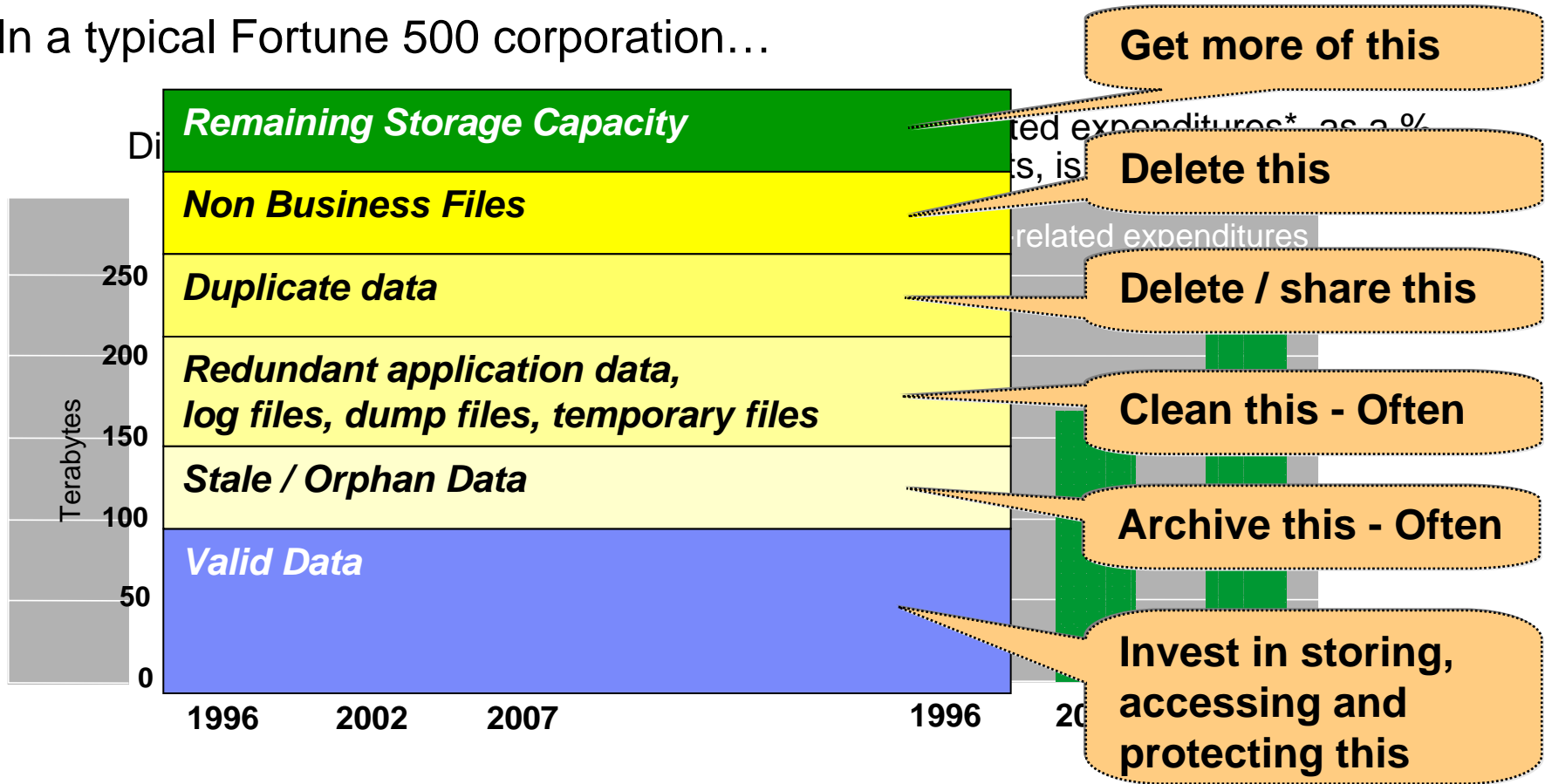
Source: International Technology Group, Sept 2003

\* hardware, software, storage networking, personnel, backup operations, recovery, security



# With an On Demand Storage Environment... Data is categorized

In a typical Fortune 500 corporation...



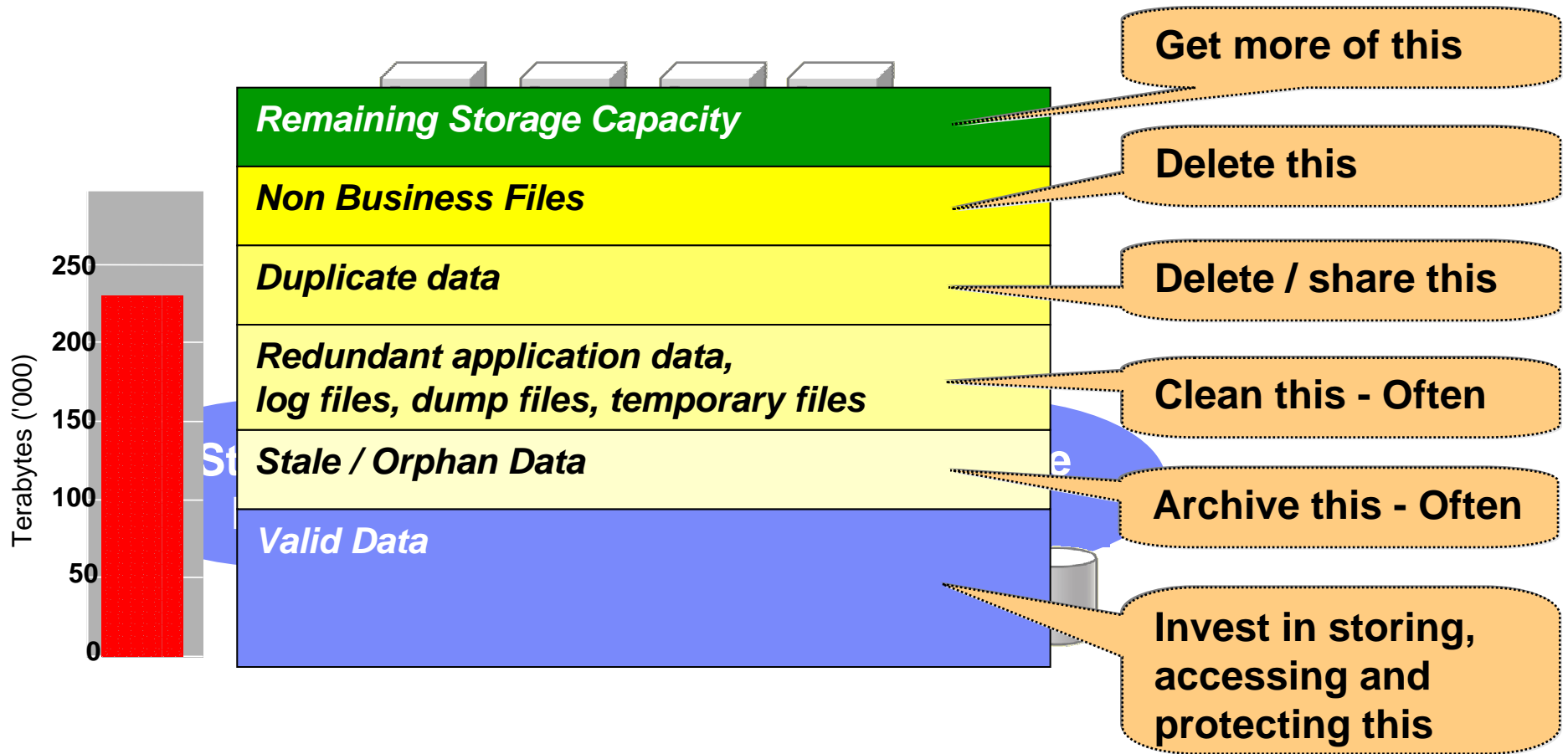
Source: International Technology Group, Sept 2003

\* hardware, software, storage networking, personnel, backup operations, recovery, security



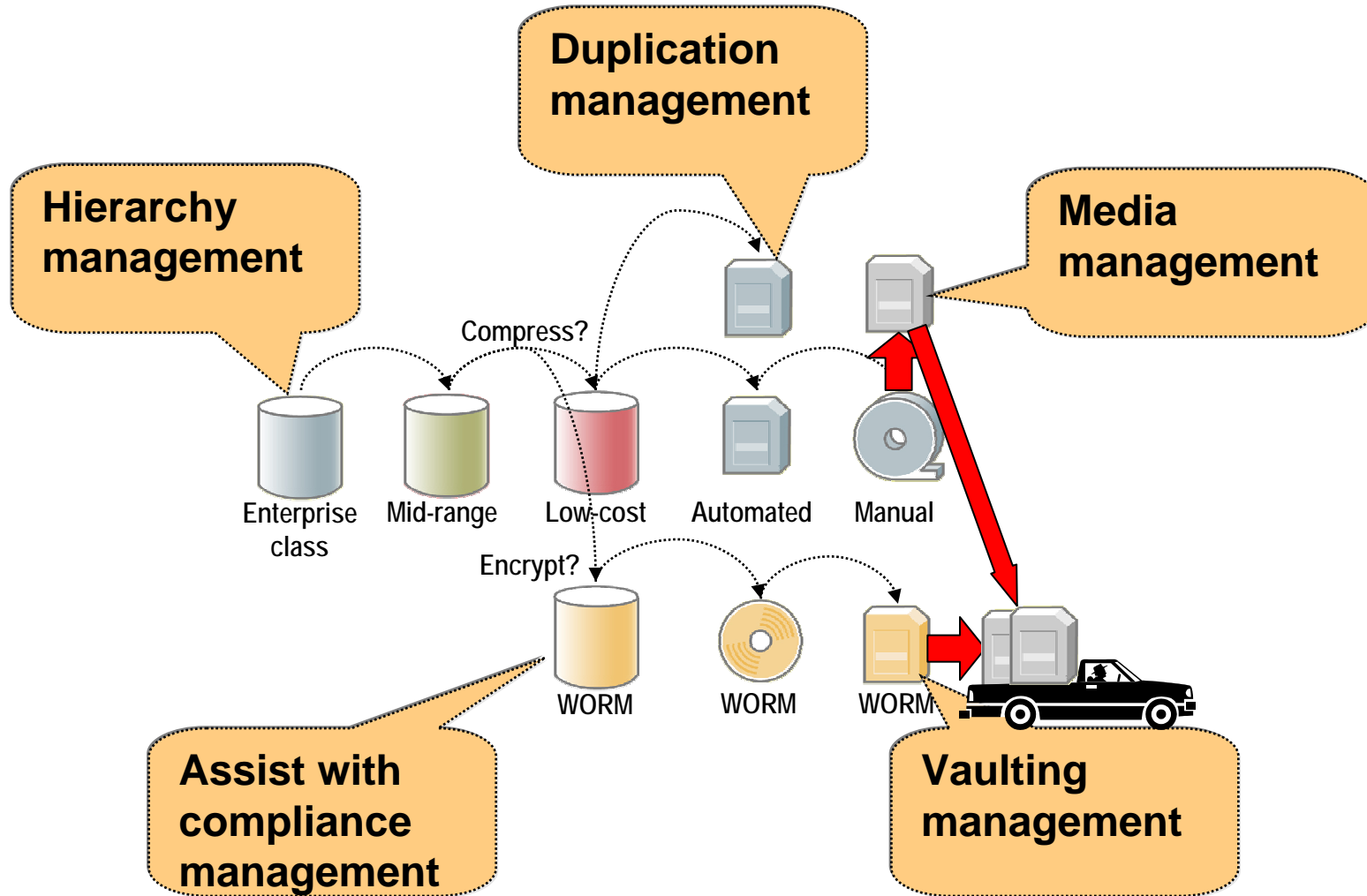
# With an On Demand Storage Environment...

Active files are pooled

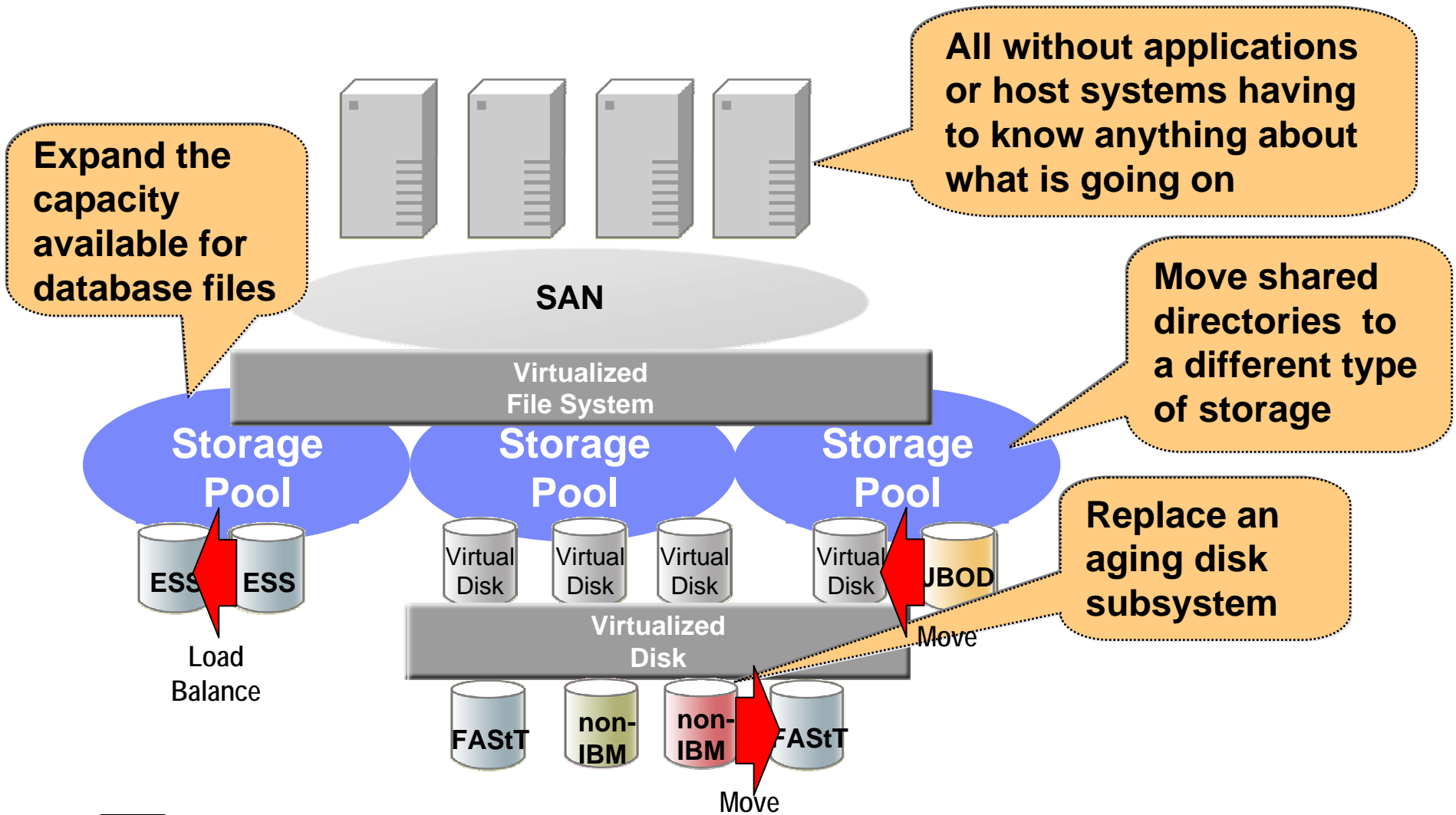


# With an On Demand Storage Environment...

Inactive files are stored in a variable-cost storage hierarchy

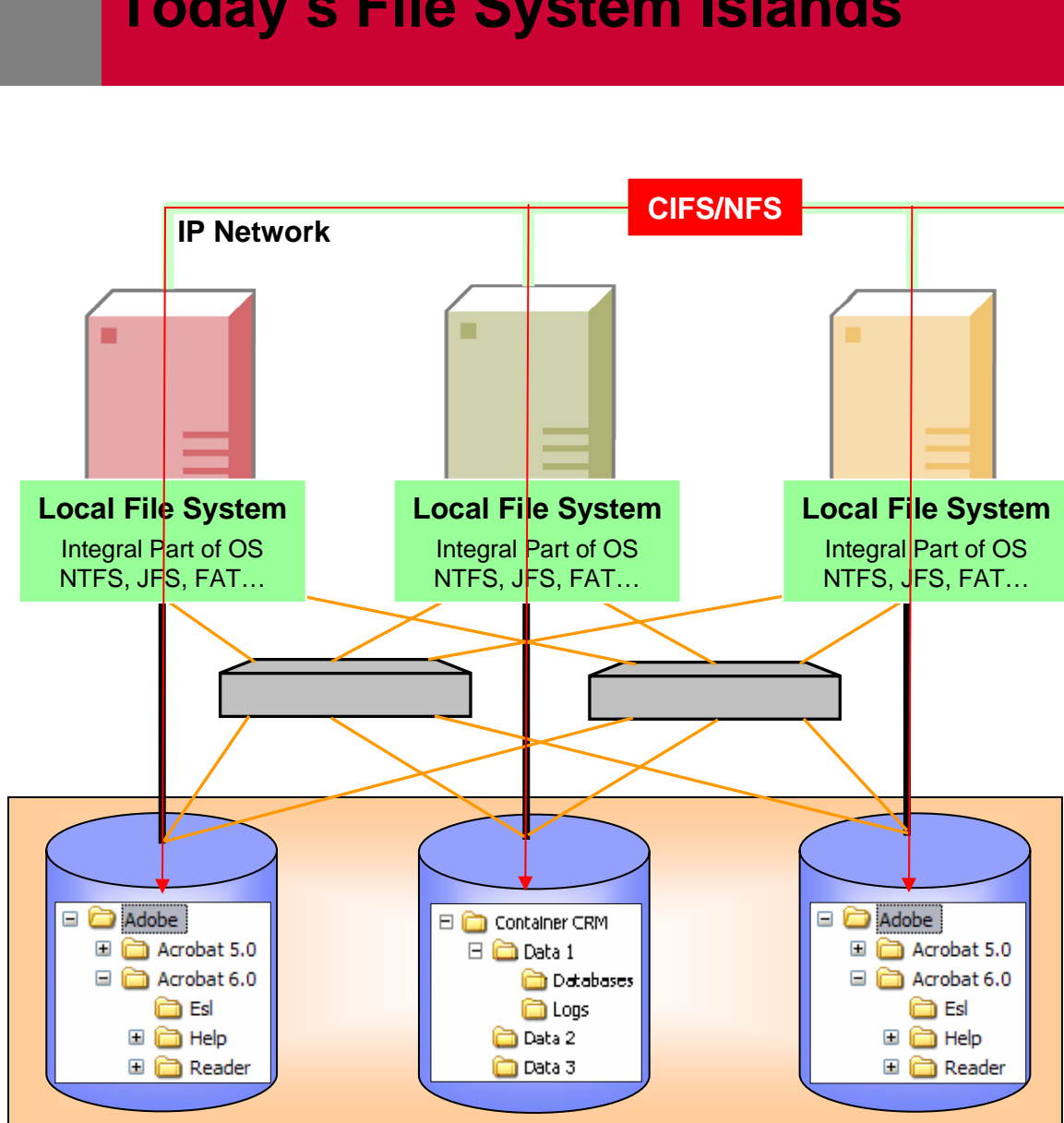


# With an On Demand Storage Environment... Infrastructure changes are non-disruptive



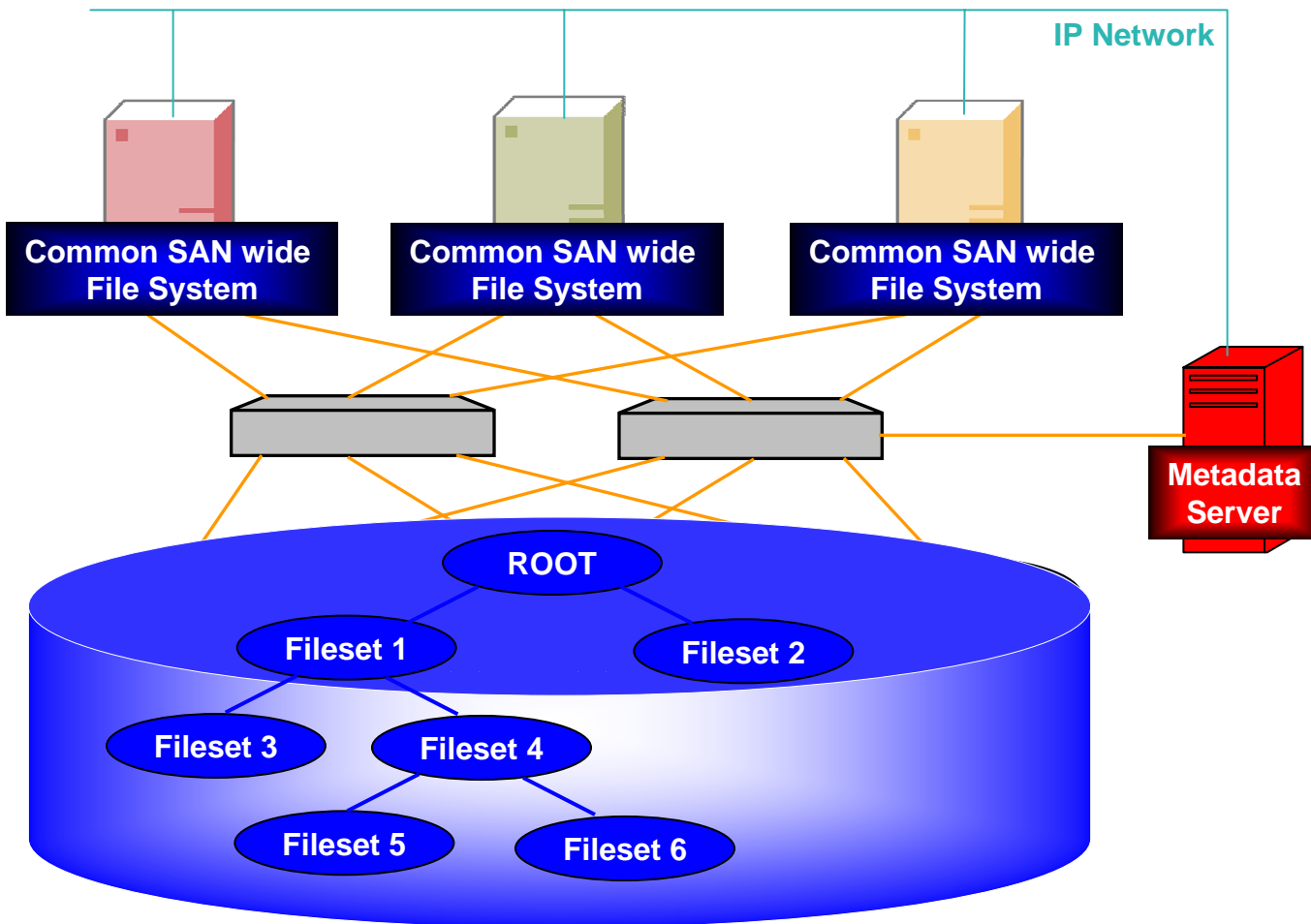


# Today's File System Islands



- ⇒ Individually managed
- ⇒ Underutilized capacity
- ⇒ Scalability issues
- ⇒ Availability issues
- ⇒ Performance issues
- ⇒ No DLM functions
- ⇒ Snapshots on box level
- ⇒ **No Global Namespace**

# The Global Namespace Solution

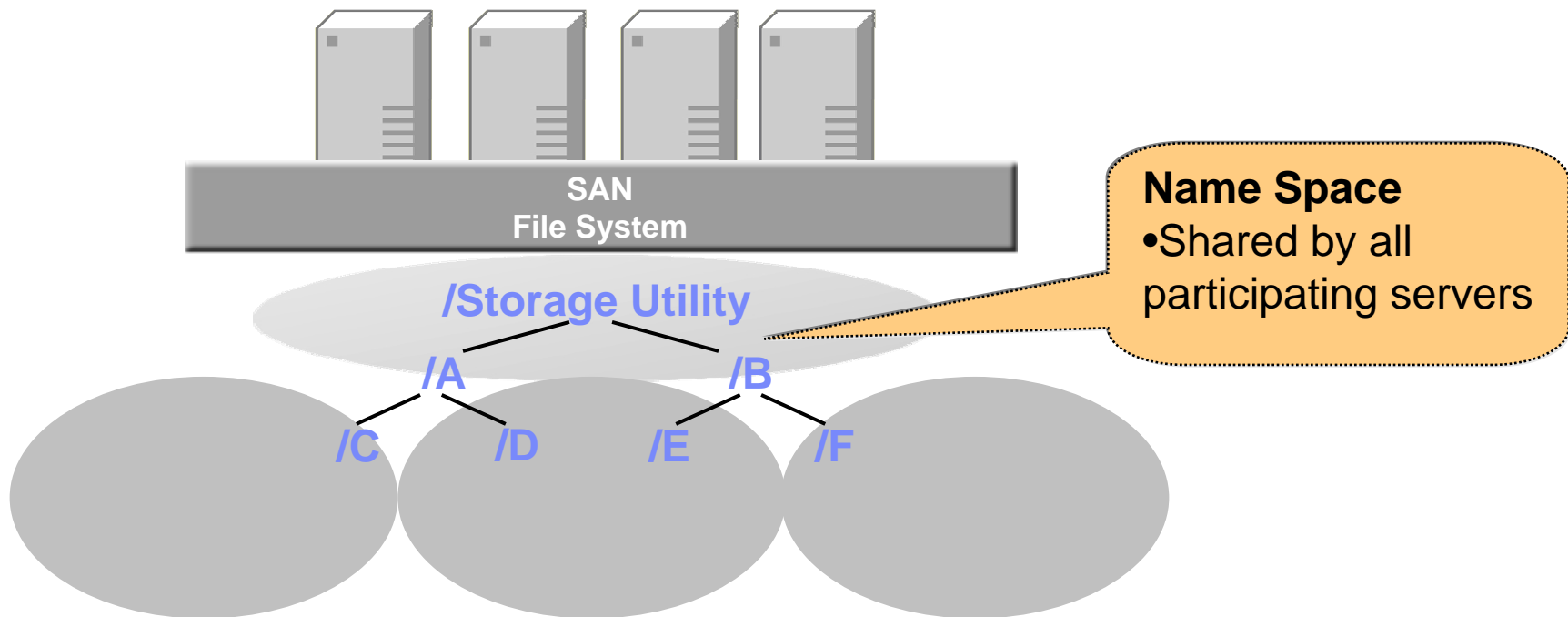


- ⇒ Shared storage
- ⇒ Single namespace
- ⇒ file sharing
- ⇒ Single point of admin
- ⇒ Infrastructure for DLM

- ⇒ All the clients
- ⇒ All the files
- ⇒ All the time!!!



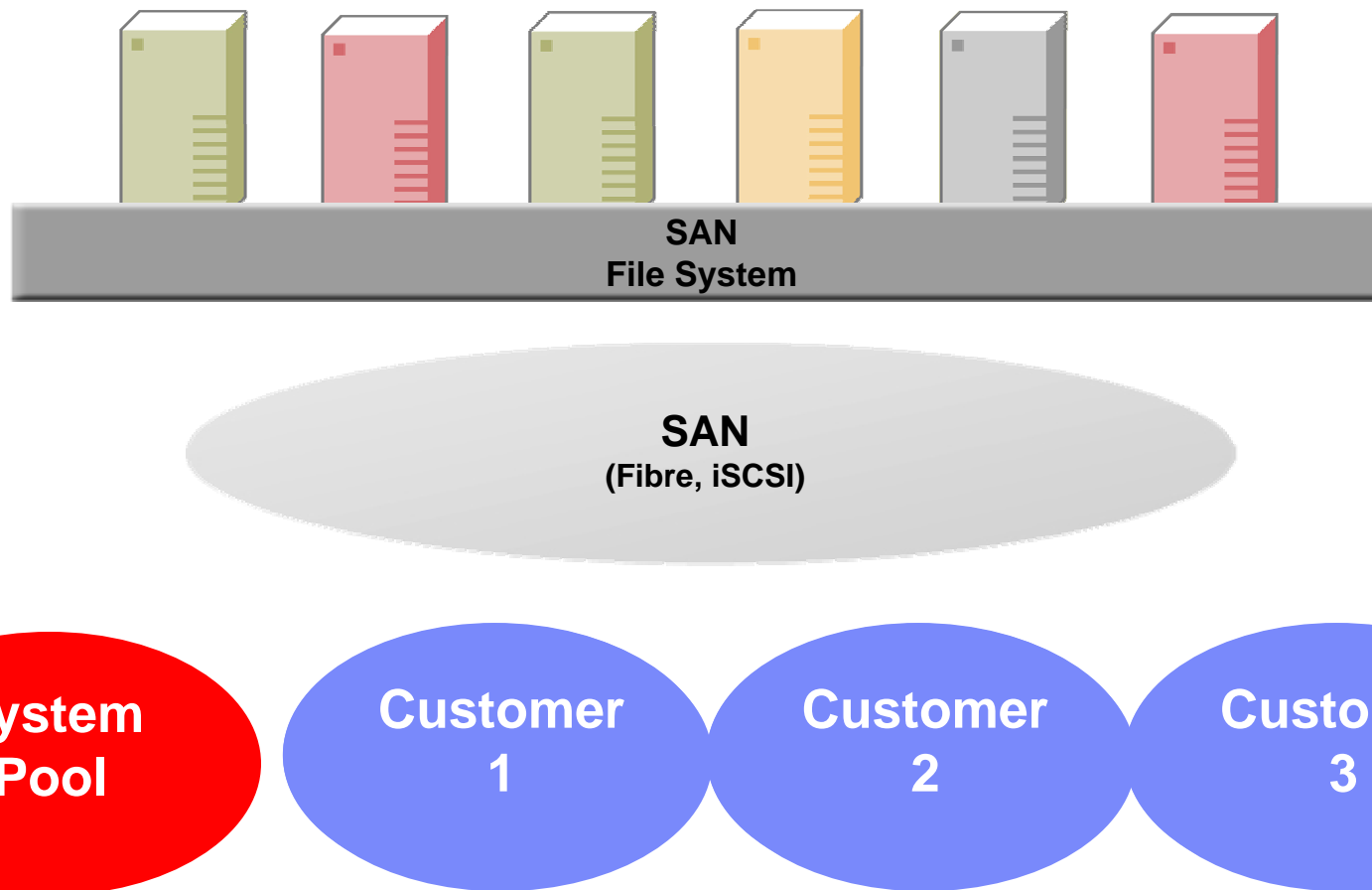
# IBM TotalStorage SAN File System



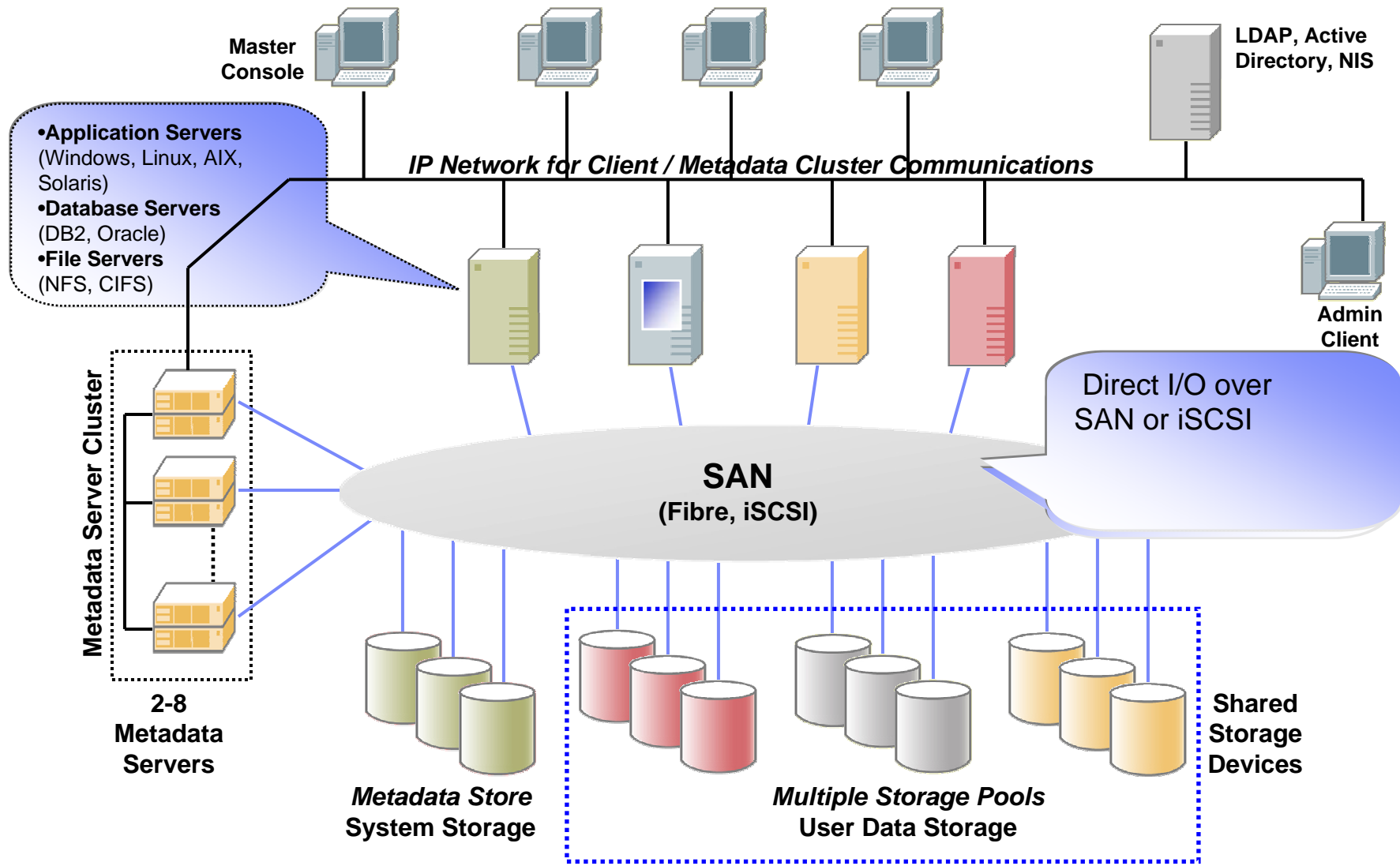
Platform	OS Version
AIX (HACMP)	5.1, 5.2, 5.3
Linux	RHE 3.0, Suse 8.0
Solaris (Cl.)	9
Windows	2000, 2003



# Storage Pools



# IBM SAN File System



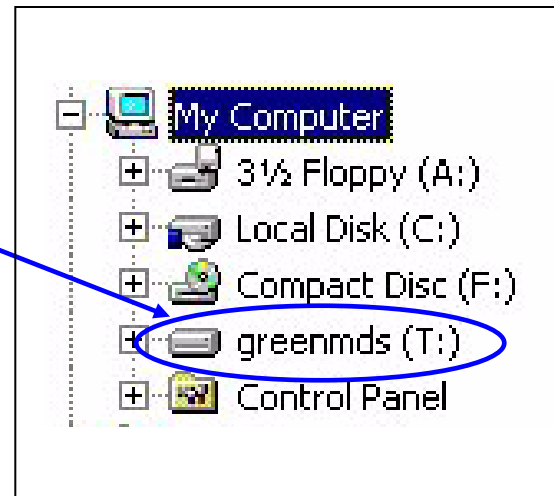
# Client View

SAN FS appears as a new drive in Windows 2000

- Supports Long file names and UNICODE
- Supports NTFS Access Control List

SAN FS appears as a new File System in UNIX / Linux

- "Visible" for native UNIX commands like df
- does not support mkfs



```
root@aix2:/# df
Filesystem 1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4      32768      23024  30%     1413   9% /
/dev/hd2     950272       8096 100%    29103  13% /usr
/dev/hd9var   32768      25640  22%     597   8% /var
/dev/hd3     65536      45452  31%     139   1% /tmp
/dev/hd1     131072     52460  60%      34   1% /home
/proc         -           -     -         -    - /proc
/dev/hd10opt  32768      24424  26%     341   5% /opt
SANFS       16728064 16154624  4%         1   1% /sanfs
root@aix2:/#
```



## The IBM SAN File System - Why use it ?

- ⇒ File Sharing – homogenous and heterogeneous
- ⇒ Policy Based Storage Automation - DLM
- ⇒ Improve Storage Utilization
- ⇒ File Level Flashcopy
- ⇒ Non disruptive Infrastructure Changes
- ⇒ Infrastructure Consolidation
- ⇒ Single Point of Management

## The IBM SAN File System - Why use it ?

- ⇒ File Sharing – homogenous and heterogeneous
- ⇒ Policy Based Storage Automation - DLM
- ⇒ Improve Storage Utilization
- ⇒ File Level Flashcopy
- ⇒ Non disruptive Infrastructure Changes
- ⇒ Infrastructure Consolidation
- ⇒ Single Point of Management



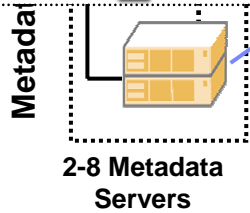
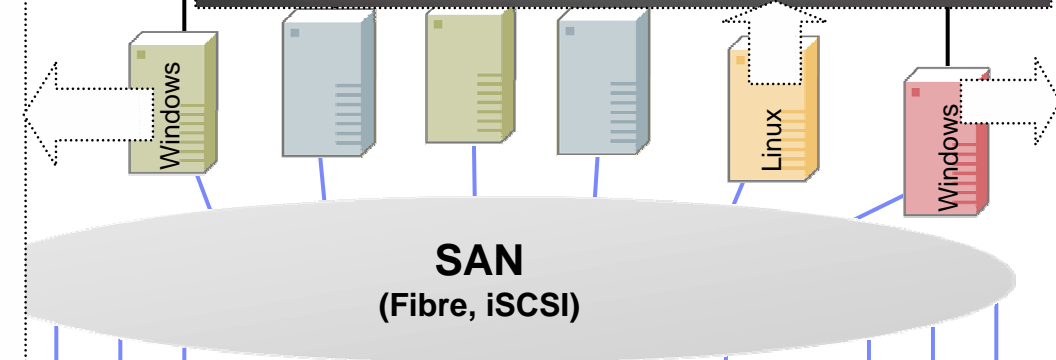
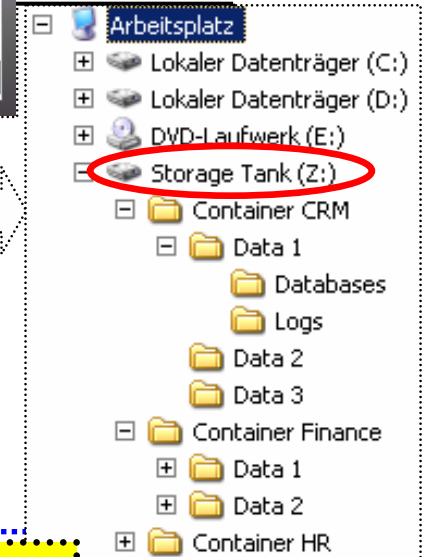
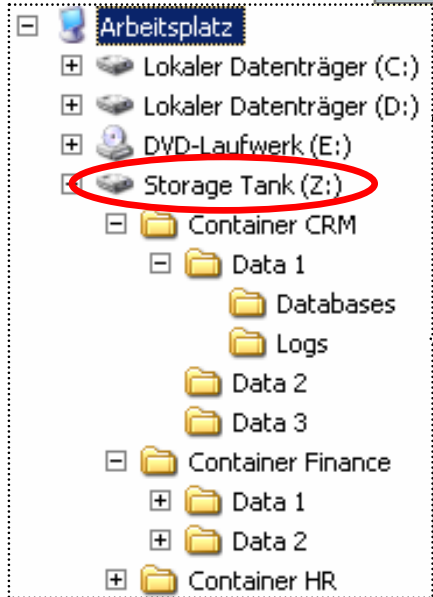
# File Sharing

```

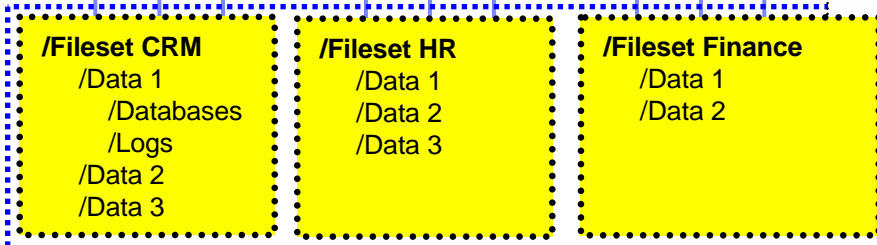
Stefan@linux:/ - Shell - Konsole
Session Edit View Bookmarks Settings Help
linux:/StorageTank #
linux:/StorageTank #
linux:/StorageTank #
linux:/StorageTank # dir
total 12
dr-x----- 1 root    root    4096 2003-05-20 18:08 .
drwxr-xr-x 22 root    root    592  2003-05-20 19:26 ..
dr-x----- 1 root    root    4096 2003-05-20 17:58 Container CRM
dr-x----- 1 root    root     0  2003-05-20 17:58 Container Finance
dr-x----- 1 root    root    4096 2003-05-20 18:08 Container HR
dr-x----- 1 root    root     0  2003-05-20 18:08 RECYCLER
dr-x----- 1 root    root     0  2003-05-20 17:53 System Volume Information
linux:/StorageTank #
linux:/StorageTank #
linux:/StorageTank #
linux:/StorageTank #
linux:/StorageTank #
linux:/StorageTank #
  
```

LDAP, Active Directory, NIS

Master Console



Metadata Store System Storage



Multiple Storage Pools User Data Storage

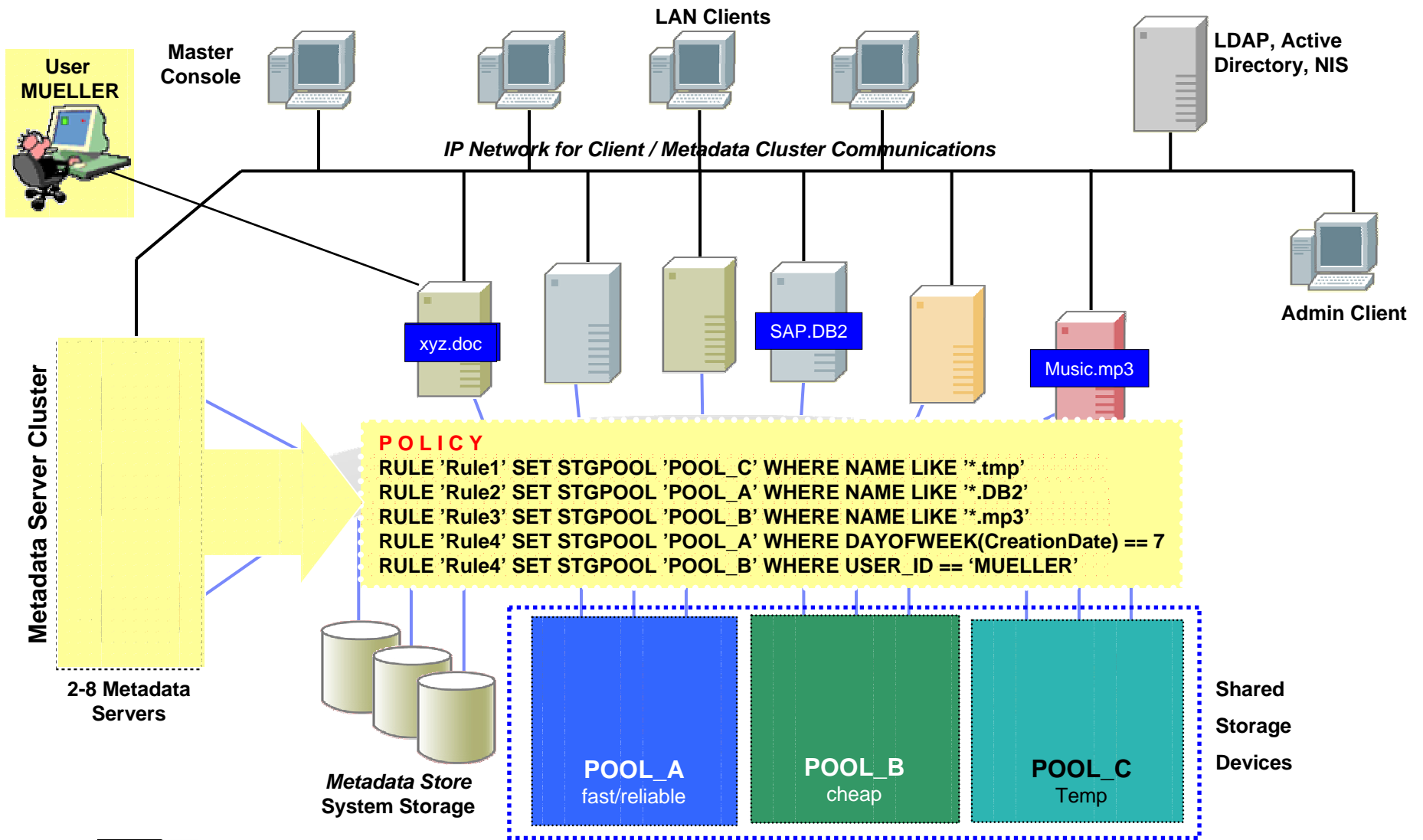
Shared Storage Devices



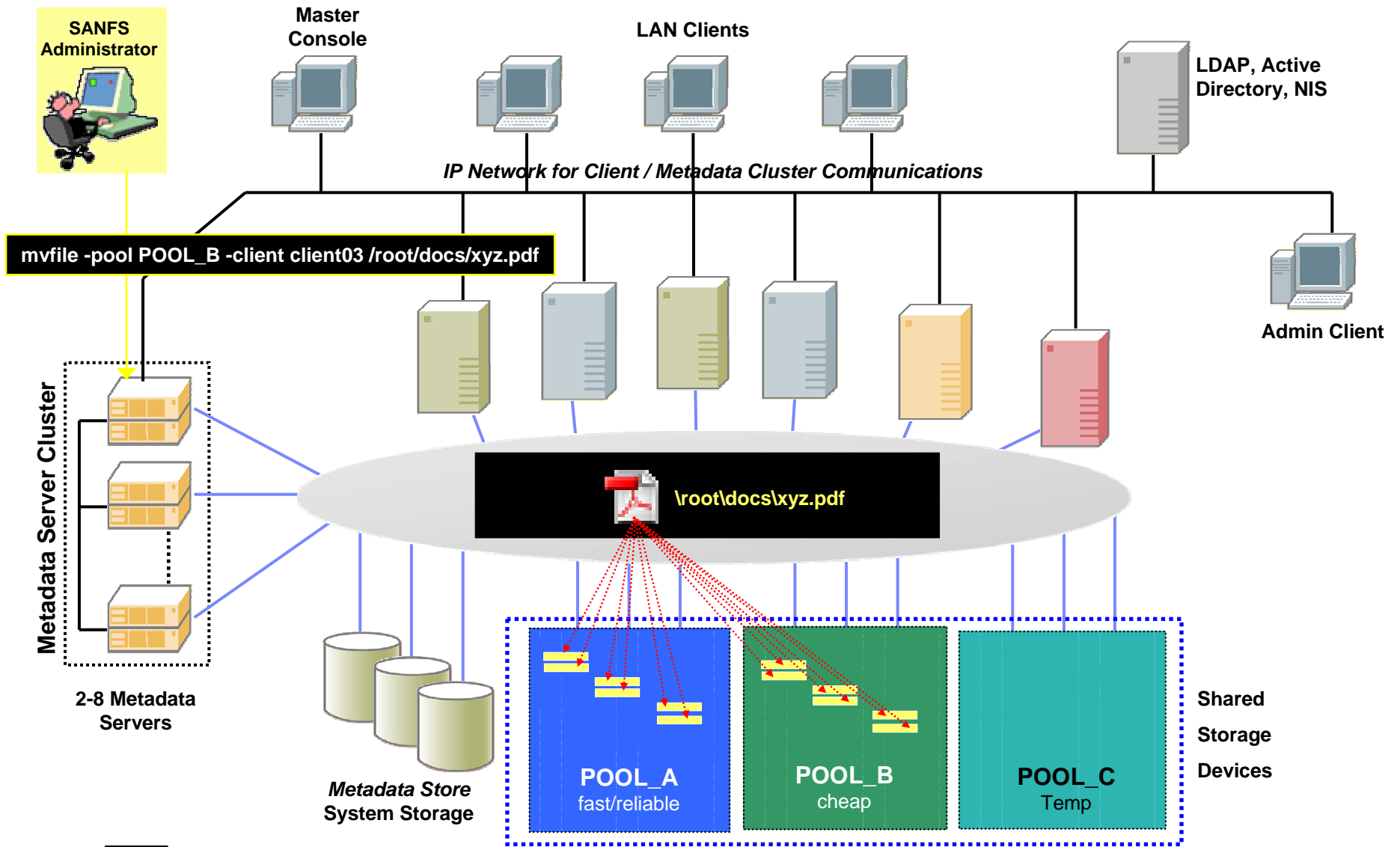
## The IBM SAN File System - Why use it ?

- ⇒ File Sharing – homogenous and heterogenous
- ⇒ Policy Based Storage Management - DLM
- ⇒ Improve Storage Utilization
- ⇒ File Level Flashcopy
- ⇒ Non disruptive Infrastructure Changes
- ⇒ Infrastructure Consolidation
- ⇒ Single Point of Management

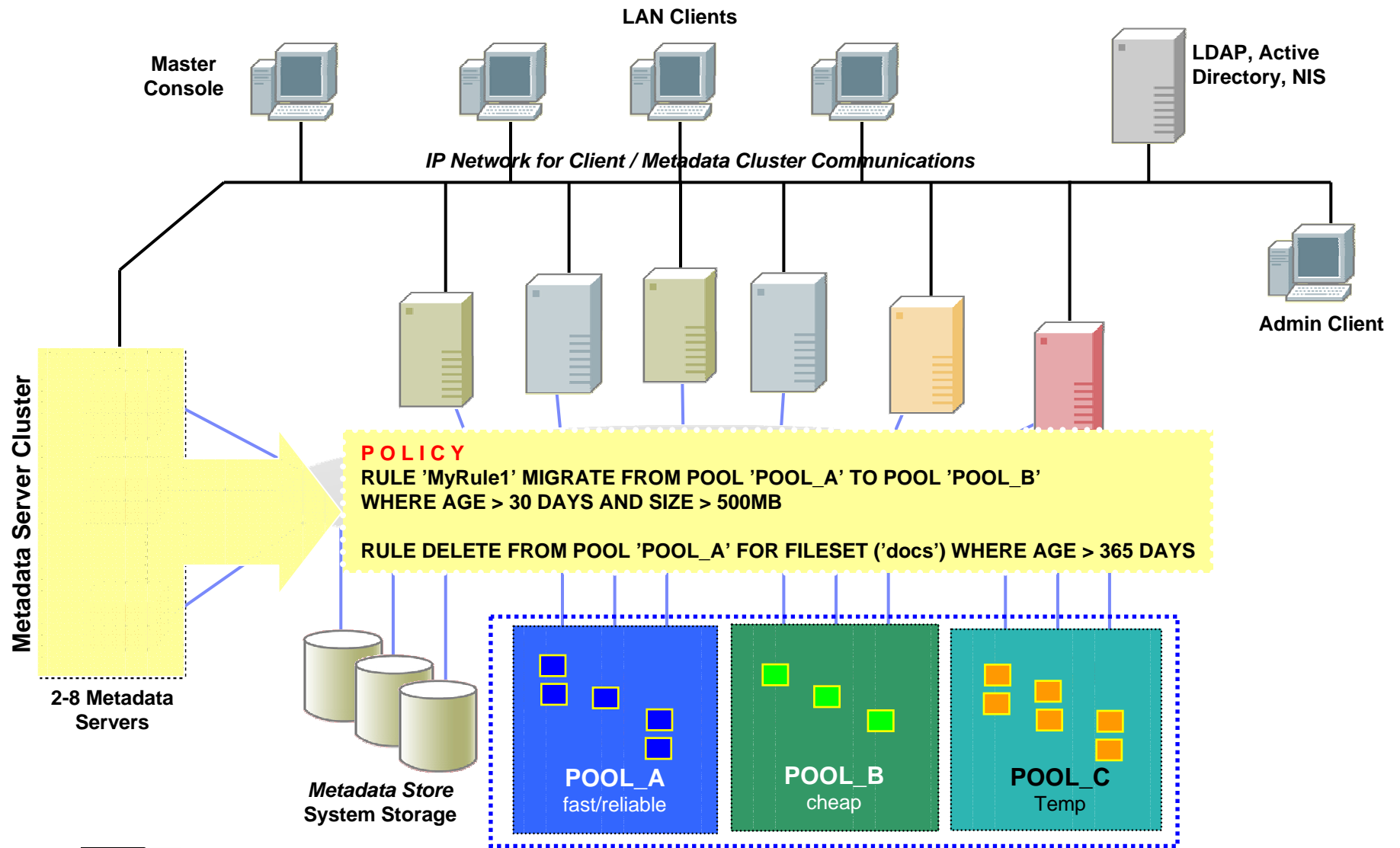
# Policy based Storage Automation – File Placement



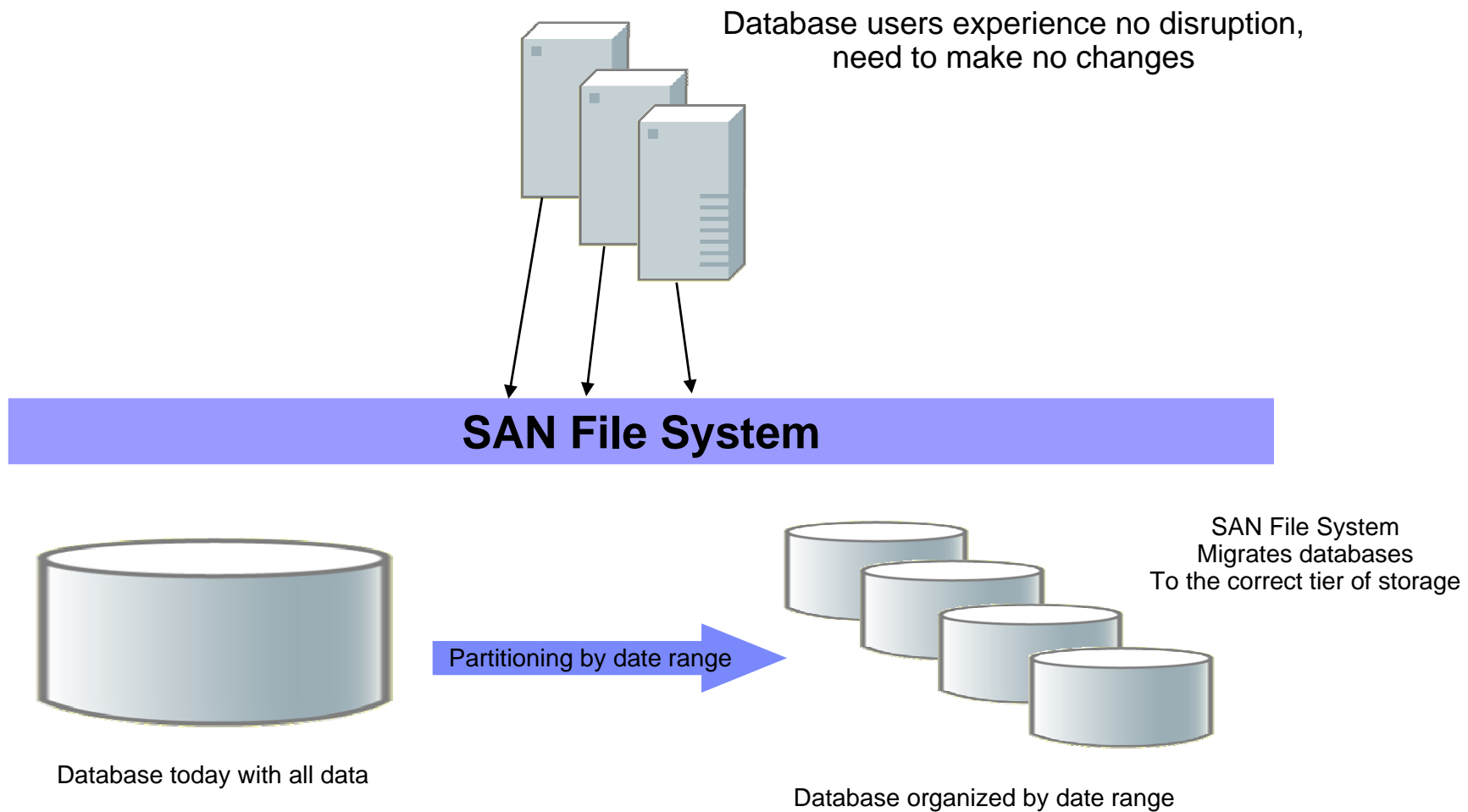
# Policy based Storage Automation – manual File Movement



# Policy based Storage Automation – Data Lifecycle Management



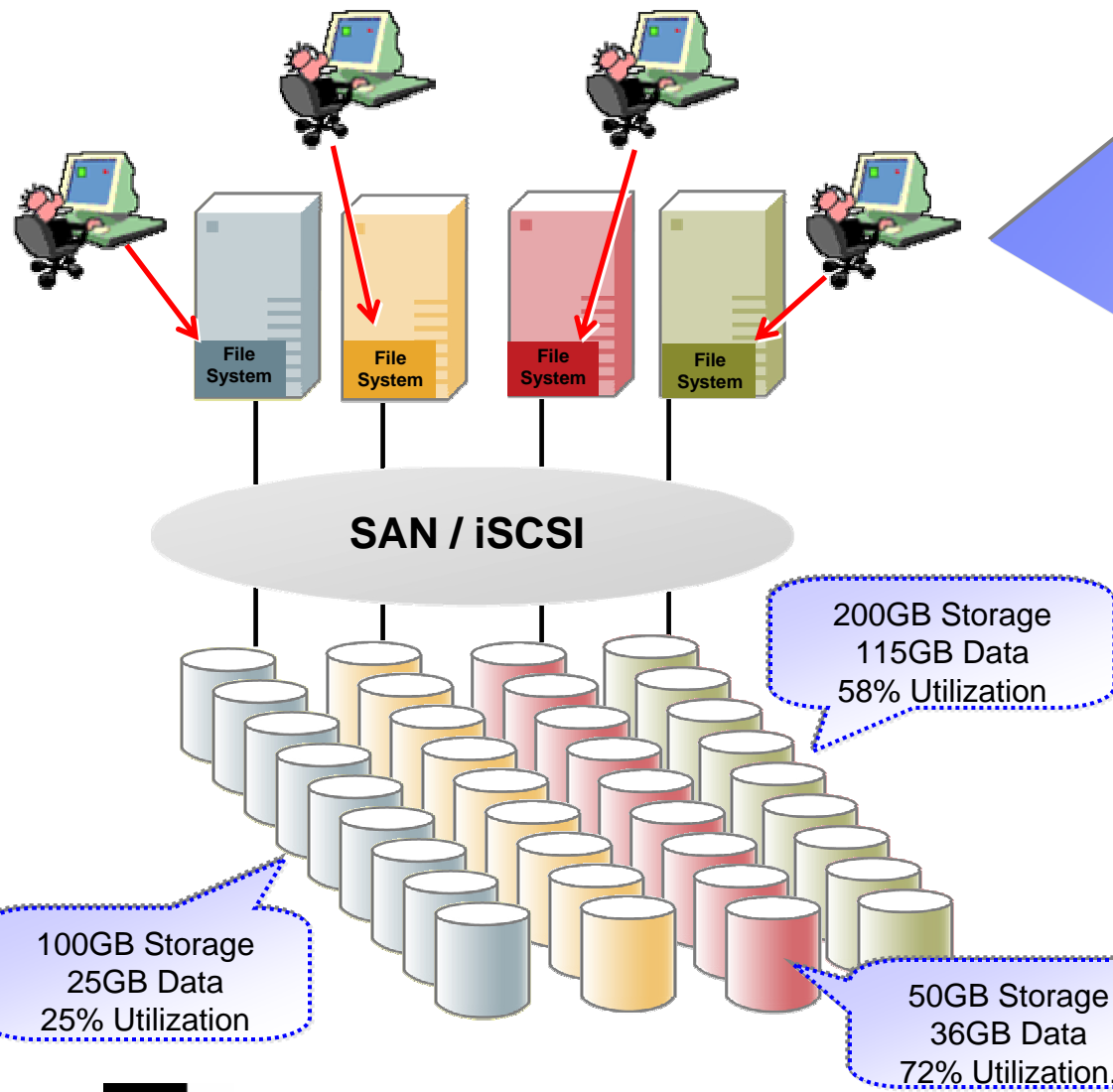
# Integration with Databases



## The IBM SAN File System - Why use it ?

- ⇒ File Sharing – homogenous and heterogenous
- ⇒ Policy Based Storage Automation - DLM
- ⇒ Improve Storage Utilization
- ⇒ File Level Flashcopy
- ⇒ Non disruptive Infrastructure Changes
- ⇒ Infrastructure Consolidation
- ⇒ Single Point of Management

# Storage Utilization Today



Each server has its own file system that must be individually configured and managed

At least one LUN per computer

Managing dozens if not 100's of LUNs

Physically dedicate disk space that may or may not be utilized

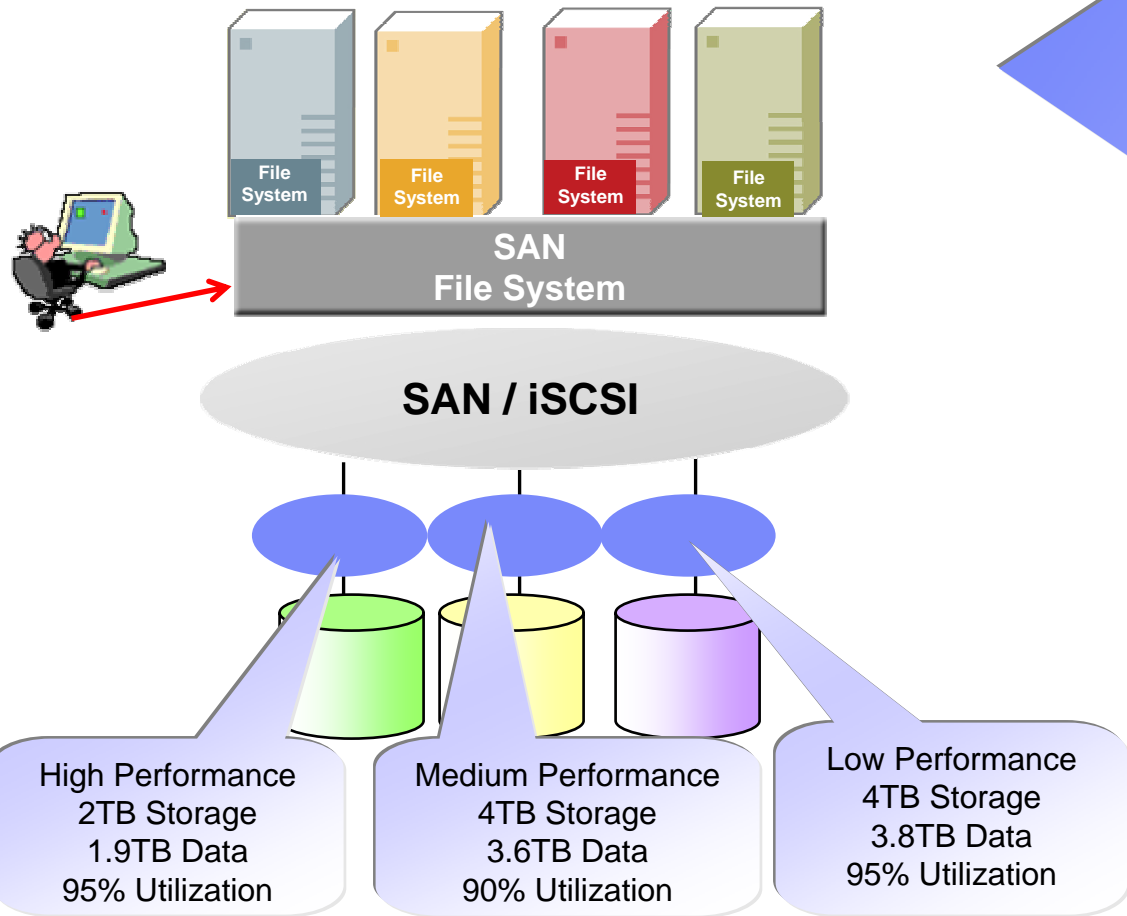
Average Utilization: 52%  
(Industry Average: 40% - 50%)

**Result...  
Poor Storage Utilization**





# Improving Storage Utilization



One file system, configured and managed centrally

Managing a few LUNs rather than dozens or hundreds

LUNs defined as storage pools based on business needs

Host over allocation of storage space enables simplified management and higher efficiencies

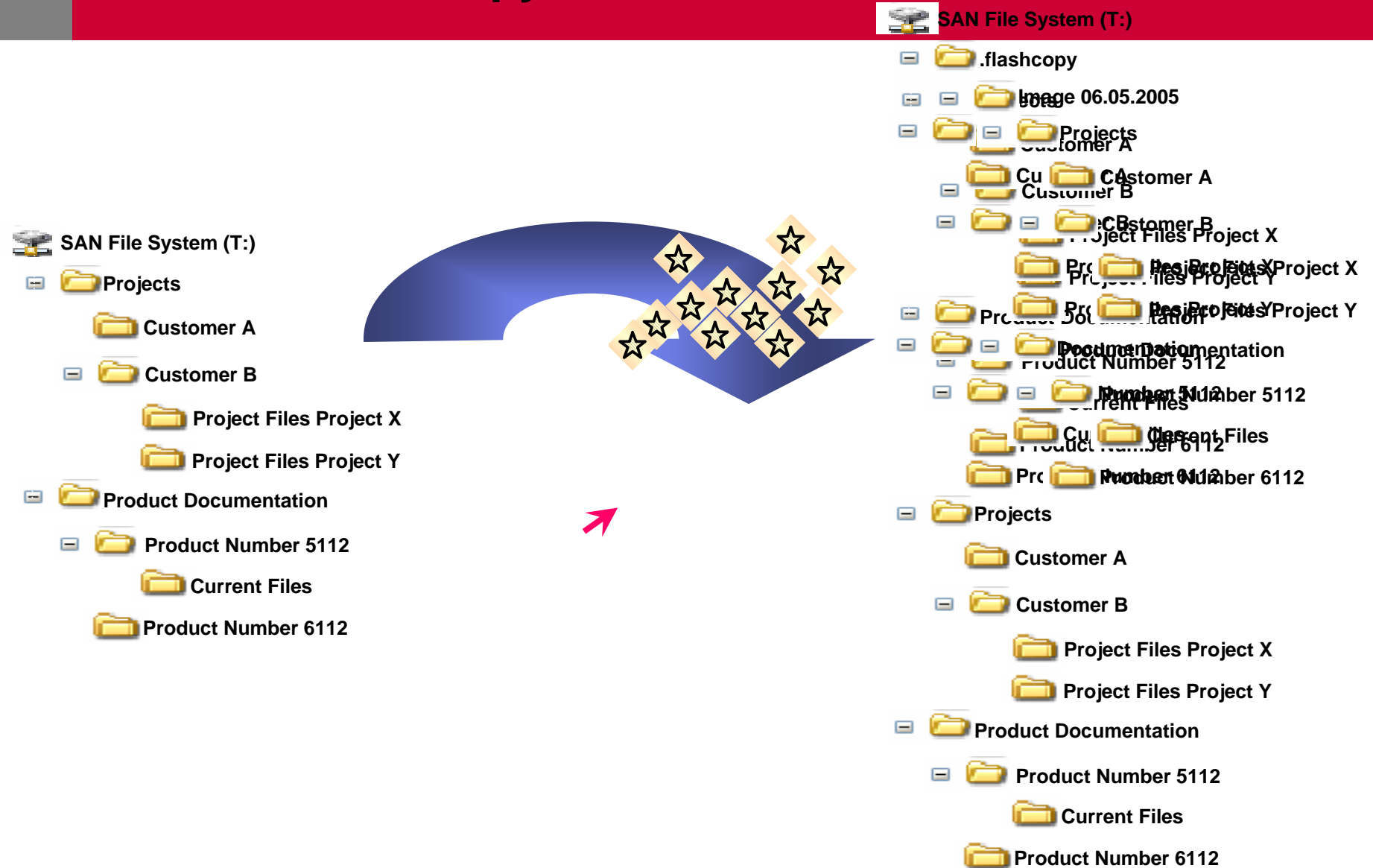
Average Utilization: 93%

**Result...  
Optimized Storage Utilization**

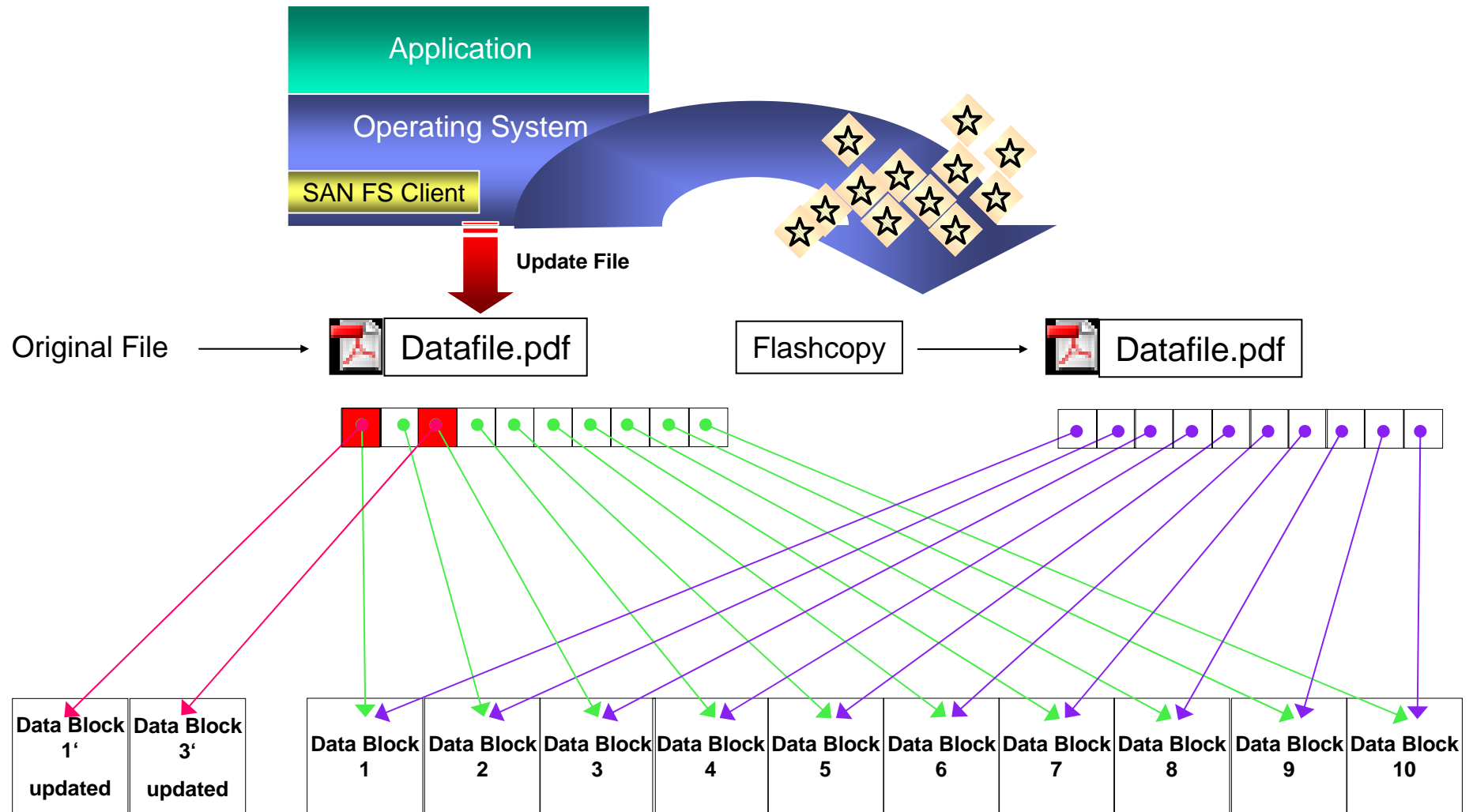
## The IBM SAN File System - Why use it ?

- ⇒ File Sharing – homogenous and heterogenous
- ⇒ Policy Based Storage Automation - DLM
- ⇒ Improve Storage Utilization
- ⇒ File Level Flashcopy
- ⇒ Non disruptive Infrastructure Changes
- ⇒ Infrastructure Consolidation
- ⇒ Single Point of Management

# File Level Flashcopy



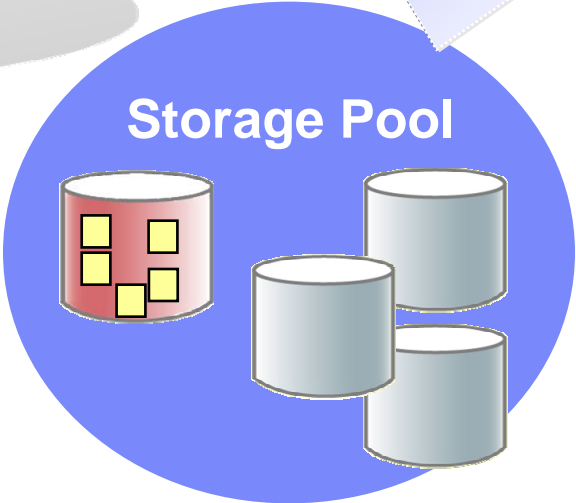
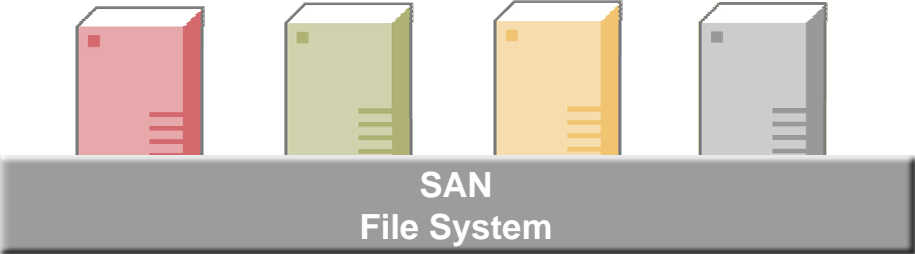
# File Level Flashcopy



## The IBM SAN File System - Why use it ?

- ⇒ File Sharing – homogenous and heterogenous
- ⇒ Policy Based Storage Automation - DLM
- ⇒ Improve Storage Utilization
- ⇒ File Level Flashcopy
- ⇒ Non disruptive Infrastructure Changes
- ⇒ Infrastructure Consolidation
- ⇒ Single Point of Management

# Non Disruptive Infrastructure Changes



**Volume Drain**

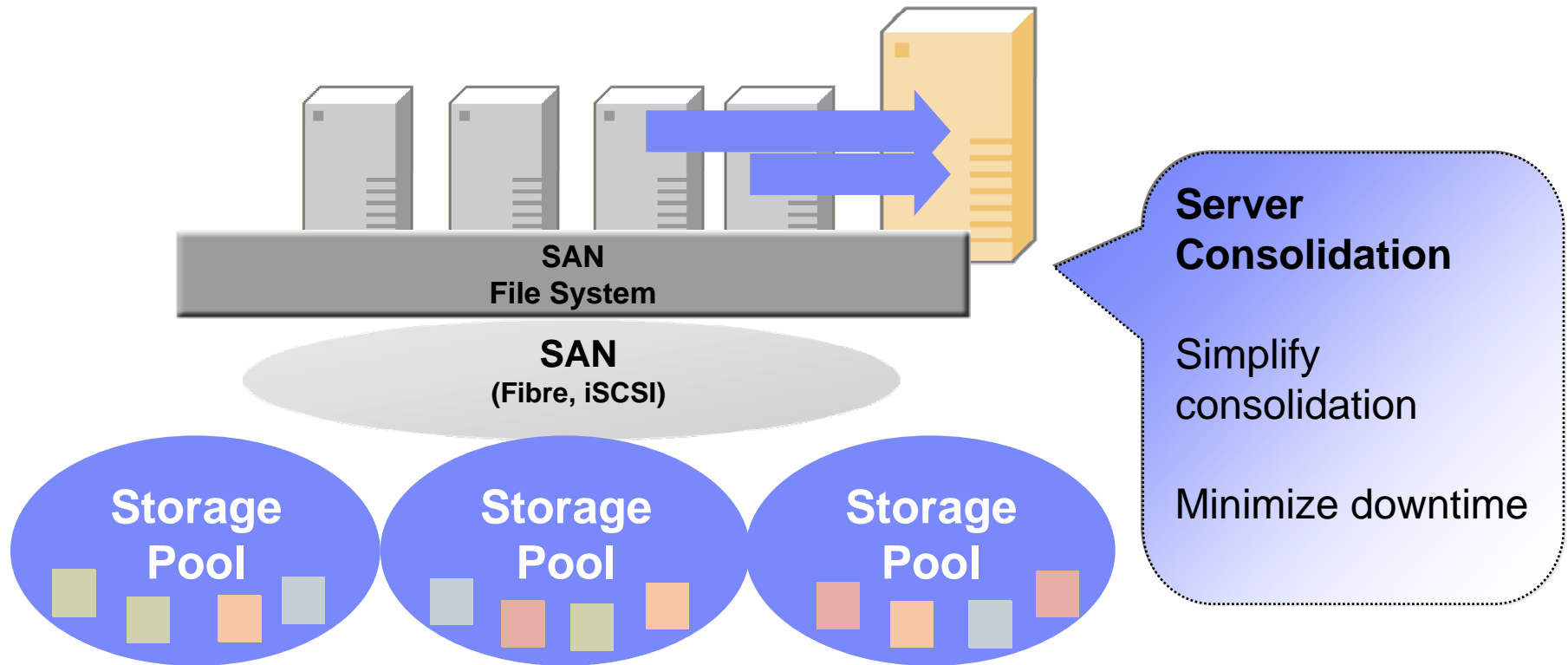
Automatically move files to like and unlike storage devices in the same storage pool

No need to stop application processing during volume drain

## The IBM SAN File System - Why use it ?

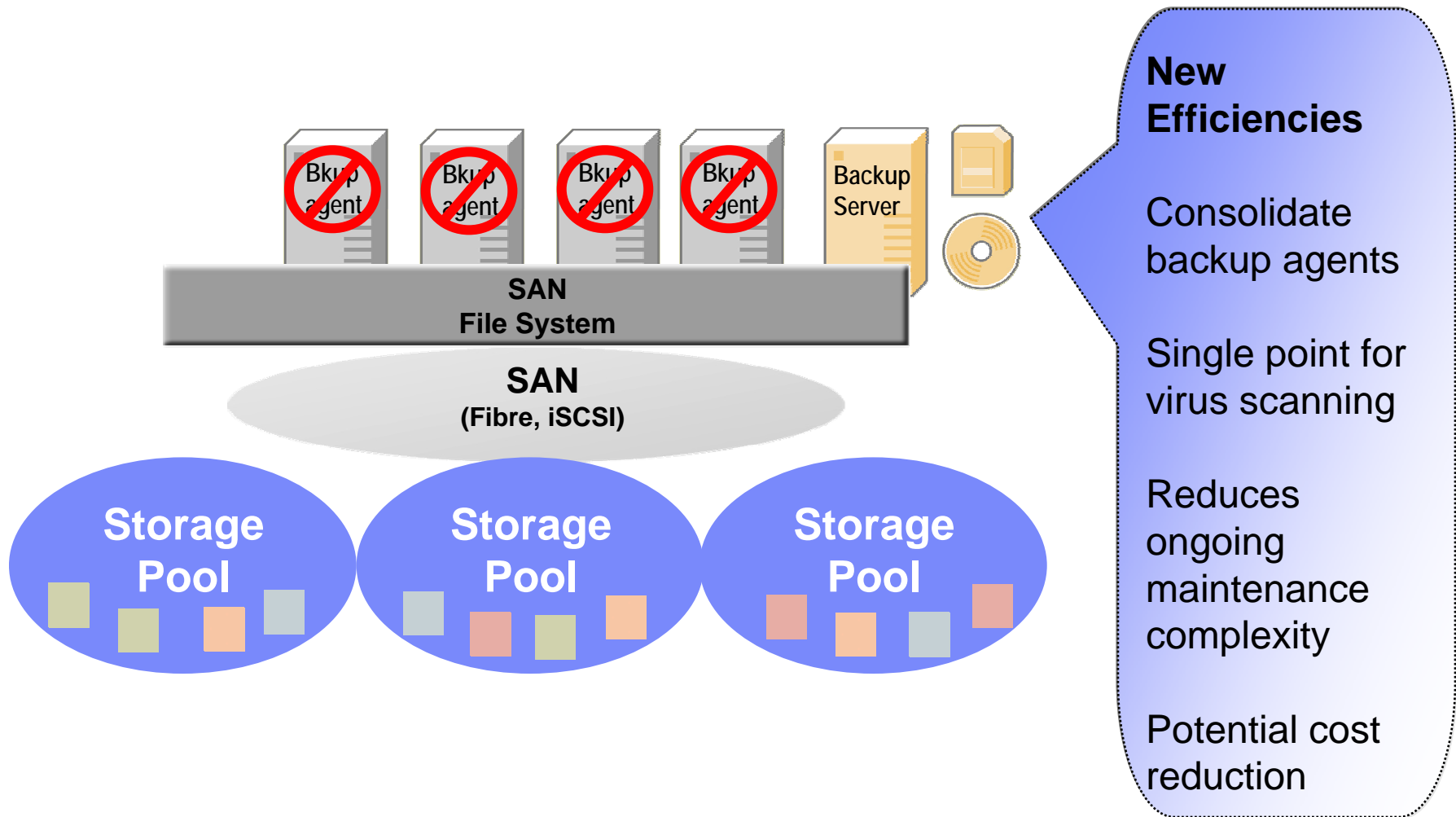
- ⇒ File Sharing – homogenous and heterogenous
- ⇒ Policy Based Storage Automation - DLM
- ⇒ Improve Storage Utilization
- ⇒ File Level Flashcopy
- ⇒ Non disruptive Infrastructure Changes
- ⇒ Infrastructure Consolidation
- ⇒ Single Point of Management

# Server Consolidation - Minimize Application Downtime





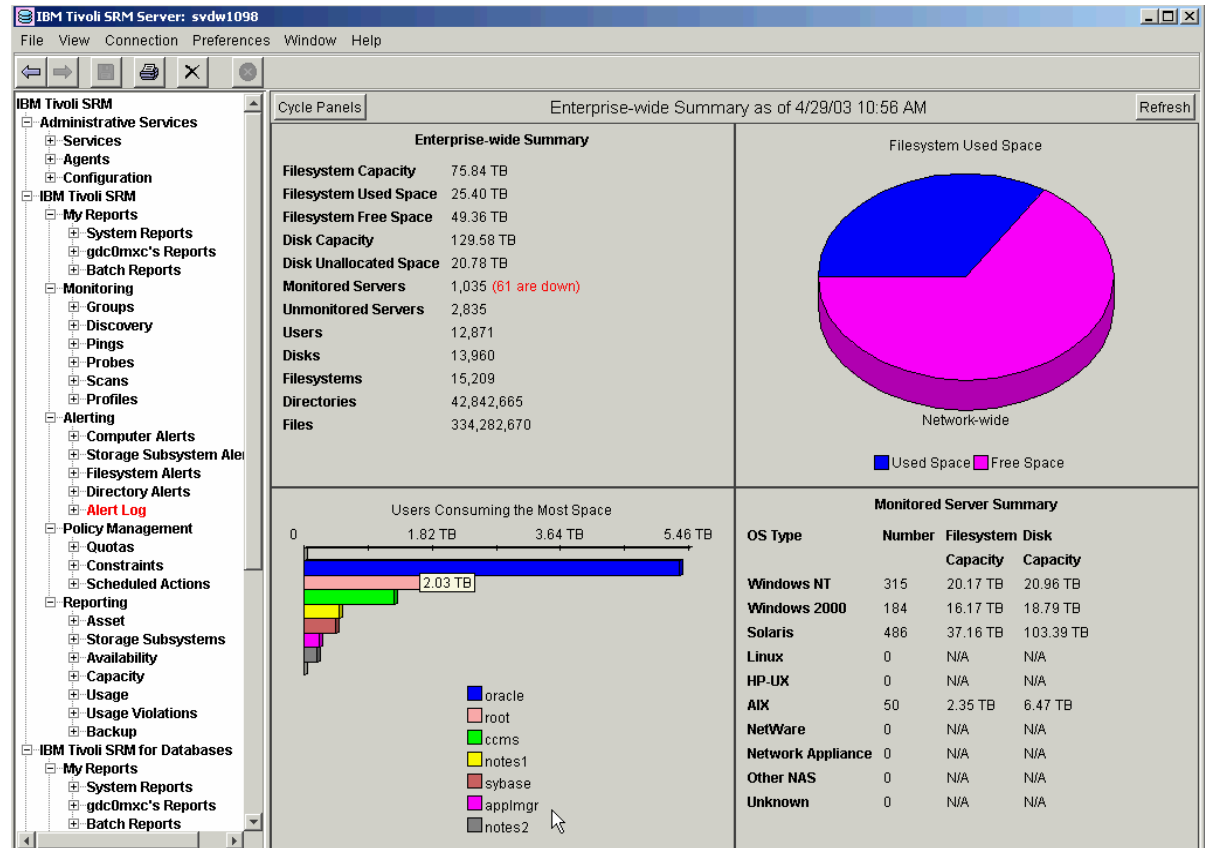
# Backup Consolidation - Simplify Operations With Centralized Services



# IBM TotalStorage Productivity Center for Data single point for accounting

## Monitor

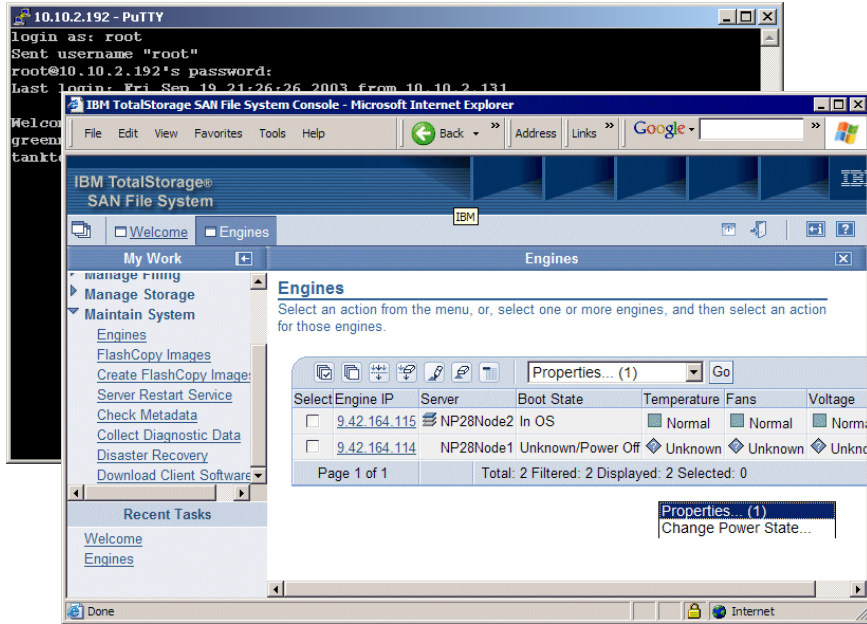
- ⇒ What are your storage assets?
- ⇒ Do your allocations match expectations?
- ⇒ What is your current utilization?
- ⇒ Do you have at-risk file systems?
- ⇒ Do you have allocated, but unused database space?



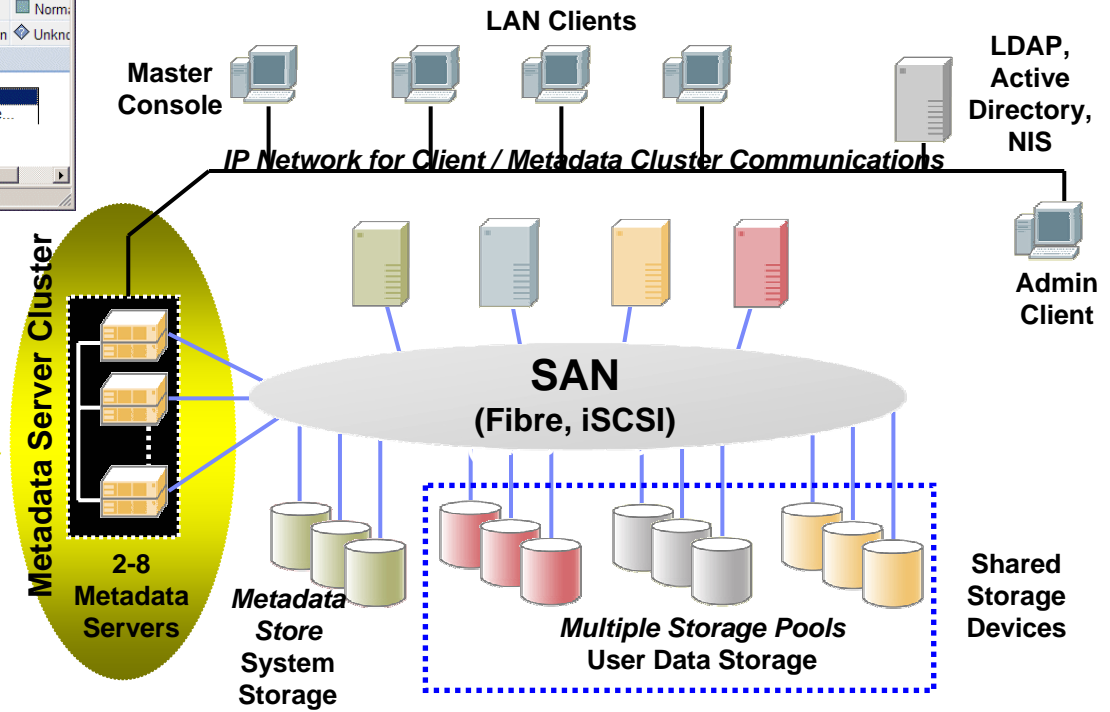
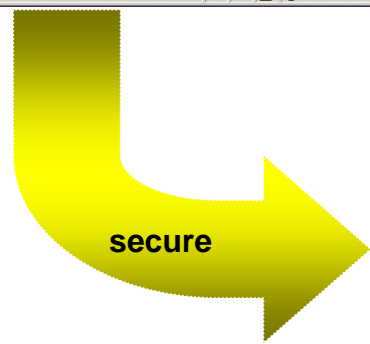
### The IBM SAN File System - Why use it ?

- ⇒ File Sharing – homogenous and heterogenous
- ⇒ Policy Based Storage Automation - DLM
- ⇒ Improve Storage Utilization
- ⇒ File Level Flashcopy
- ⇒ Non disruptive Infrastructure Changes
- ⇒ Infrastructure Consolidation
- ⇒ Single Point of Management

# Single Point of Management

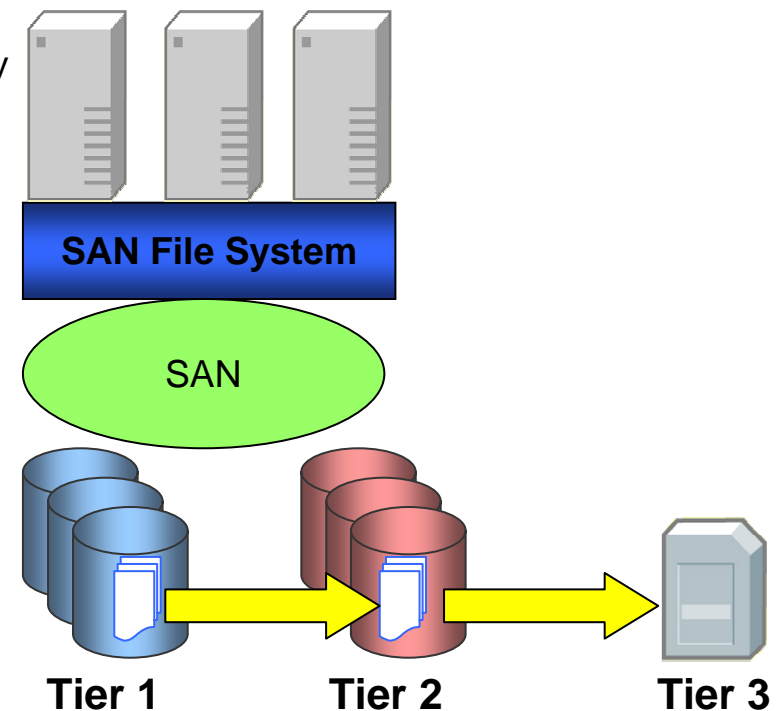


- ⇒ performed via
  - SAN FS Console (Web-based)
  - Command Line Interface - tanktool
- ⇒ SSH / SSL
- ⇒ Authentication against LDAP



# IBM TotalStorage SAN File System V3.1 (target GA 1Q2006)

- **Enhance ILM policy language to support additional selection criteria**
  - File create time (age)
  - File modification time (time since last modified)
  - File name
- **Extend policy based tiered storage management to include tape**
  - Integrated with TSM for migrating data to tape or secondary disk storage
  - SAN File System controls migration policy to move files to TSM
  - Treats TSM tape pool or disk storage as a virtual storage pool
  - Metadata of offline objects (files moved to TSM) is maintained in SAN File System namespace
  - SAN File System policy control over deletion of offline files
  - Support manual operations on offline files: list, delete
  - Support for AIX and Linux archive clients only (can archive Windows files using these clients)



## Business-Case (standard – no ILM)

- 100 Server
- 100 TB Storage capacity
- File-System usage 50%
- All data stored on highend storage devices (DS8000)
  - purchase costs: 20€/ GB
  - Migration costs: 8€/ GB
  - Depreciation period 36 months
  
- Purchase costs = 2,0 Mio €
- Migration costs = 0,8 Mio €
  
- → 2,8 Mio € / 50TB used cap. / 36 month =  
1,56 €/ used TB / month

# Business-Case (with SAN-FS)

- 100 Servers (100 processors)
- 100 TB Storage capacity
- File-System usage 90%
- 30% of the data is stored on highend storage devices (DS8000)
- 30% of the data is stored on “medium class” storage devices (DS4500)
- 40% of the data is stored on low-end storage devices (DS4100)
  - HW purchase costs: 20€/ GB (DS8000), 10€/ GB (DS4500), 5€/ GB (DS4100)
  - No Migration costs:
  - Depreciation period 36 months
- HW Purchase costs =  $100.000 * (0,3*20 + 0,3*12 + 0,4*5) = 1,16 \text{ Mio } \text{€}$
- SW-Costs: 40.000€(MetaDataServer),  $100 * 2.000 \text{ €(Clients) + 100.000 €SW-Maintenance for 3 years} = 0,34 \text{ Mio } \text{€}$
- →  $1,5 \text{ Mio } \text{€} / 90\text{TB used cap.} / 36 \text{ month} = 0,46 \text{ €} / \text{ used TB} / \text{ month}$   
(1,56 € / used TB / month in a non SAN-FS environment)

## Business-Case 2 (with SAN-FS)

- **200 Servers with 2 processors each (400 processors in total)**
- **100 TB Storage capacity**
- **File-System usage 90%**
- **30% of the data is stored on highend storage devices (DS8000)**
- **30% of the data is stored on “medium class” storage devices (DS4500)**
- **40% of the data is stored on low-end storage devices (DS4100)**
  - HW purchase costs: 20€/ GB (DS8000), 10€/ GB (DS4500), 5€/ GB (DS4100)
  - No Migration costs:
  - Depreciation period 36 months
- **HW Purchase costs = 100.000 \* (0,3\*20 + 0,3\*12 + 0,4\*5) = 1,16 Mio €**
- **SW-Costs: 80.000€(4 System-MetaDataServer), 400 \* 2.000 €(Clients) + 350.000 €SW-Maintenance for 3 years = 1,23 Mio €**
- → **2,39 Mio € / 90TB used cap. / 36 month = 0,74 €/ used TB / month**  
(1,56 €/ used TB / month in a non SAN-FS environment)



# What about compliance ?

- **Where do I have to store my “compliance” data ?**
- **Who manages retention policies ?**
- **What are the legal requirements ?**

# International Laws

Cohasset Associates, Inc.

## International Laws

### Background

- Very few laws that have media or storage specific requirements.
  - France may be the only one with a media specific “guideline” requiring optical disk storage.
- Tax revenue laws of Germany (GDPdU and GoBS) and Japan (Electronic Ledger Storage Law) specifically, as well as other European countries are similar to IRS Revenue Procedures.

### Storage-related Requirements

- Most laws and regulations have storage-related requirements similar to most U.S. laws
  - Protect the integrity of the record from unauthorized alteration and deletion.
  - Preserve the content, context and, where possible, the structure of the record, including linkages between associated records (a data transaction and the associated document-based information).
  - Make the records available or accessible.
  - Ensure that the records are “processable” and reproducible in human-readable format



**FRITZ & MACZIOL : INFOMA®**

## Some regulations tell you **how** to store data:

“(ii) The electronic storage media must:

(A) Preserve the records exclusively in a non-rewriteable, non-erasable format; .”<sup>2</sup>

<sup>2</sup> SEC 17a-4 (f) (2)

## While Some Regulations tell you **what** to do (or not to do)

### **“SEC. 802. CRIMINAL PENALTIES FOR ALTERING DOCUMENTS.**

(a) IN GENERAL.—Chapter 73 of title 18, United States Code, is amended by adding at the end the following:

#### **“§ 1519. Destruction, alteration, or falsification of records in Federal investigations and bankruptcy**

“Whoever knowingly alters, destroys, mutilates, conceals, covers up, falsifies, or makes a false entry in any record, document, or tangible object with the intent to impede, obstruct, or influence the investigation or proper administration of any matter within the jurisdiction of any department or agency of the United States or any case filed under title 11, or in relation to or contemplation of any such matter or case, shall be fined under this title, imprisoned not more than 20 years, or both.”<sup>1</sup>

<sup>1</sup> Sarbanes-Oxley Act of 2002

# IBM Tivoli Storage Manager for Data Retention

- **Interfaces with ECM (Enterprise Content Management) applications through TSM-API**
- **Offers nearly the same capabilities like SAN-FS**
  - Policy based data management
  - Incorporated storage-hierarchy (HSM)
  - Planned integration with SAN-FS (TSM-API support from SAN-FS)
- **Data Retention Protection**
  - Data will not be deleted until the retention criteria for the object is satisfied
- **Usually Retention period is controlled by ECM-application (or Records Manager)**



# IBM TotalStorage Data Retention 550 (DR550)

- A comprehensive offering designed to provide non-rewriteable and non-erasable policy based storage management
- Autonomic policy based data migration
- Attaches to IP Network via secure access API
- Many redundant components for high availability
- Integrated Solution



# Summary

- DR550 is **the** solution for data archival in a regulated environment
- **SAN-FS lowers TCO and provide significant benefits for all file-system data**
  - Increase storage utilization (from 40-50% up to more than 90%)
  - Non-disruptive infrastructure-changes (expansion and shrinking of storage pools)
  - Policy based, tiered Storage Automation (better than DFHSM)
  - Automatic data migration (even if datasets are in use)
  - Space efficient FlashCopy for all data (about 10% capacity for FlashCopies)
  - Shorter Backup-windows and faster Recovery with FlashCopy.
  - Infrastructure Consolidation (Backup, Virus-Scanning, Accounting etc.)
  - Convergence of SAN and NAS
  - Single point of Management

**Why don't you start saving money today ?**

