



H-RAIN: An Architecture for Future Proofing Digital Archives

Andres Rodriguez

781-890-8353

andres@archivas.com

NASA/IEEE MSST 2004

12th NASA Goddard/21st IEEE Conference on
Mass Storage Systems & Technologies

The Inn and Conference Center

University of Maryland University College

Adelphi MD USA

April 13-16, 2004



Agenda

- ▶ Company Introduction
- ▶ Overview of Fixed Content and Requirements
- ▶ H-RAIN: An Architecture for Future Proofing Archives
- ▶ Use Case

Company Introduction

▶ History for an Online Archive

- New York Times archives

▶ Mission:

Make the software that supports the largest online archives in the world

▶ Founded in May 2003

▶ Beta Release in Q1 2004

▶ Awarded Membership in NASA and US Navy Small Business Innovation Program

▶ Experienced Team

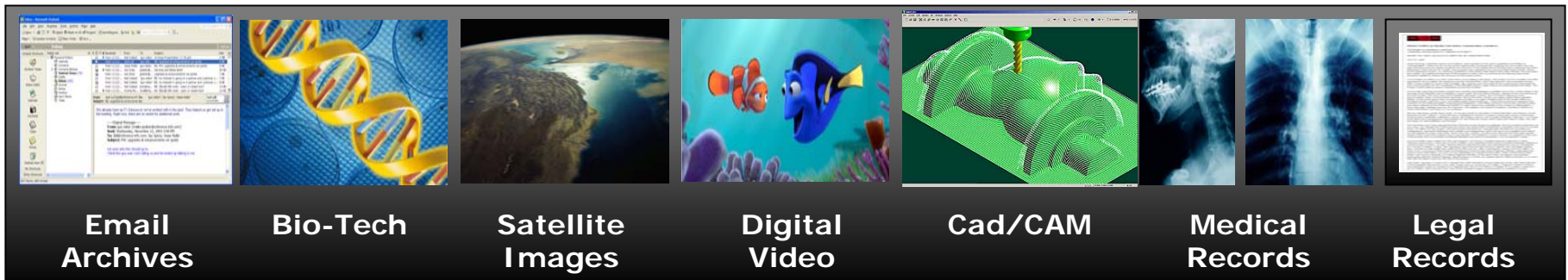
- 17 Engineers
- Seasoned Mgmt Team from NY Times, Macromedia, EMC, Compaq StorageWorks, Storability Software



Today's Archive Storage Requirements

▶ Reference Information (Fixed Content)

- Must be Write Once - Read Many: WORM



▶ Scalability

- Quantity of files: hold thousands of terabytes in a single repository
- Size of files: support for typically very large files

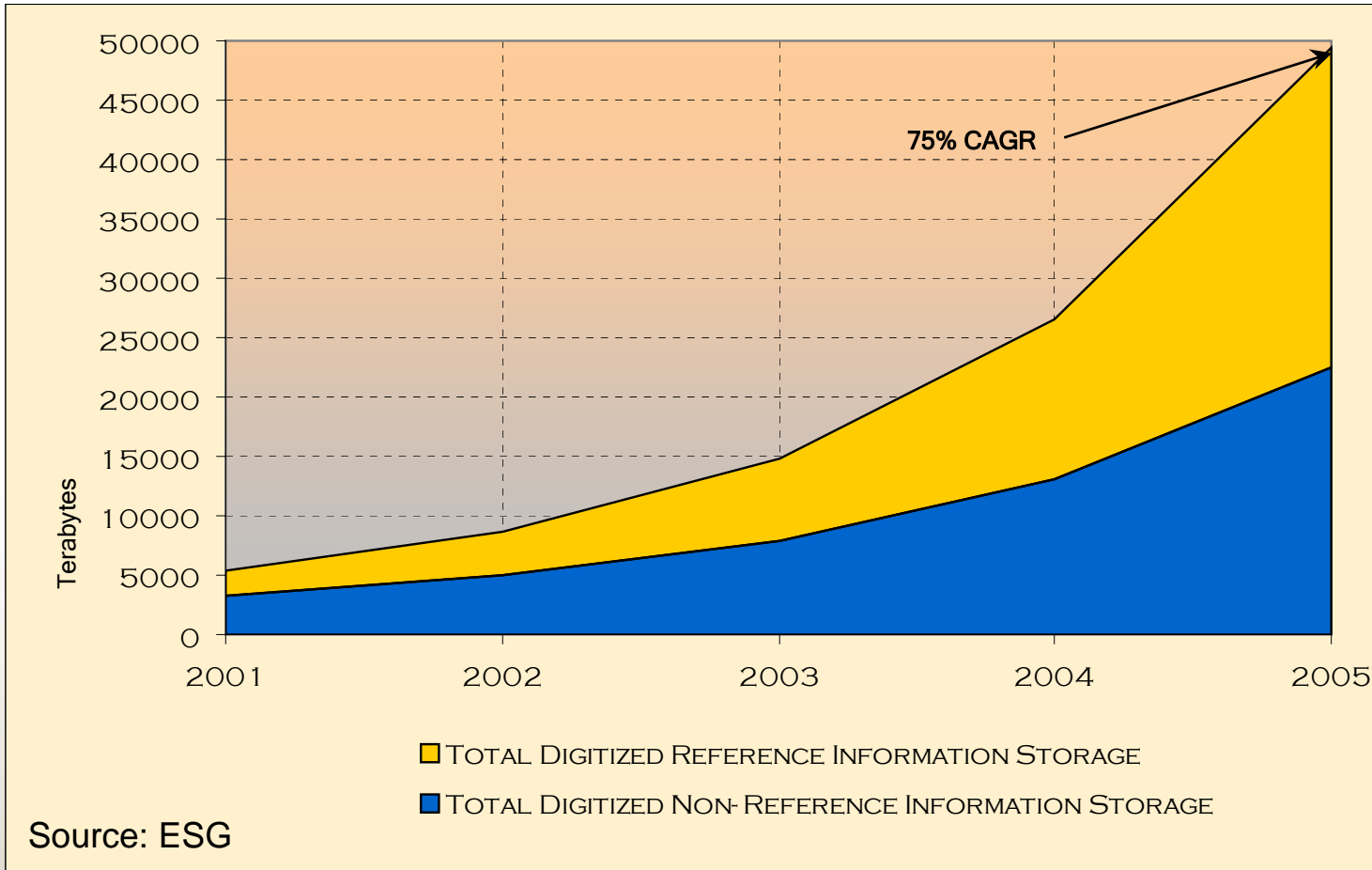
▶ Reliability

- Be able to withstand simultaneous points of failure within the archive
- Withstand site disasters

▶ Accessibility

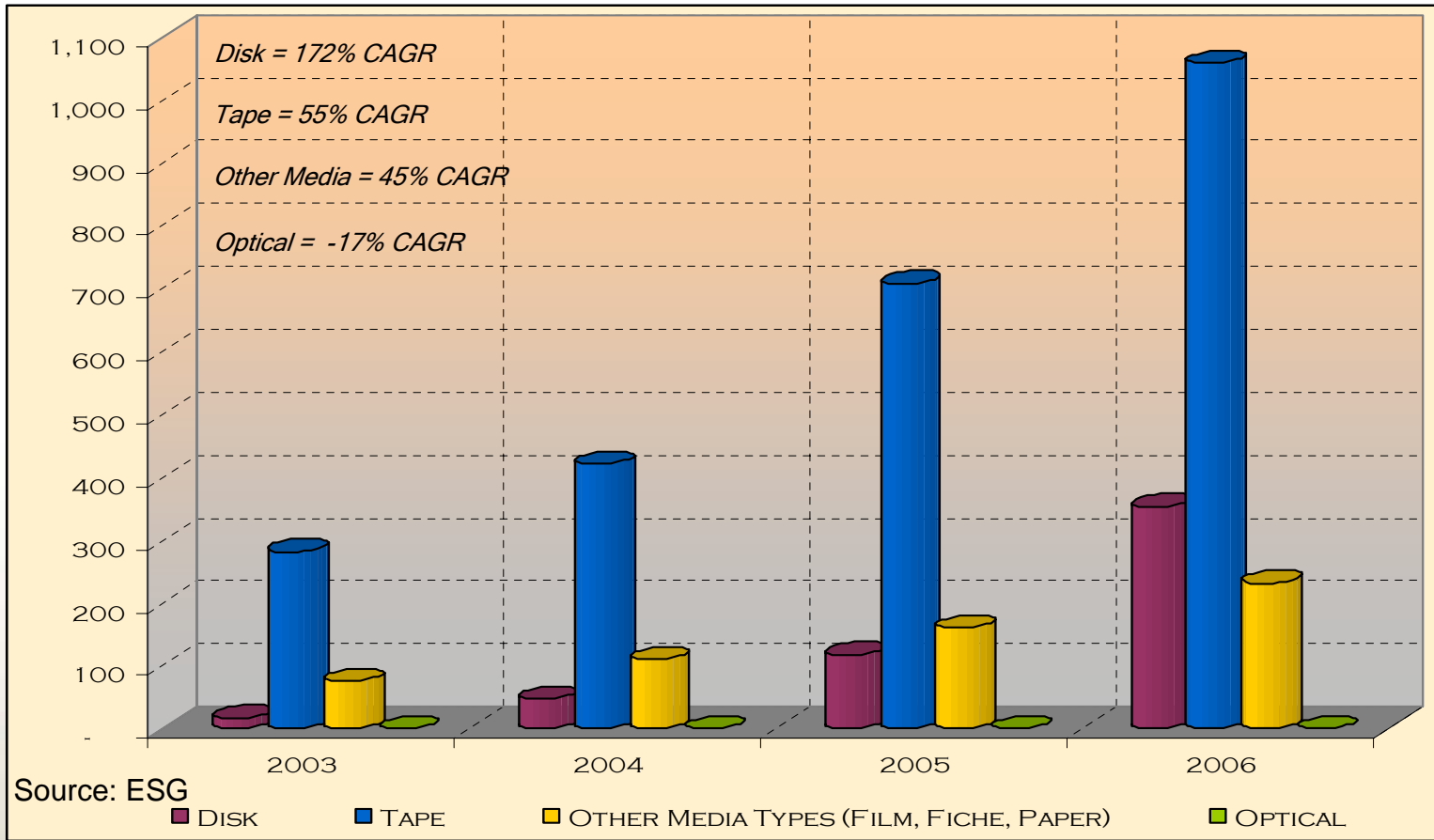
- Support and provision files to multiple applications from the same archive
- Access files using standard and open methods

Reference Information Capacity Growth



Reference Information will surpass all other information types by the end of 2004

Compliant Records Capacity by Media Type

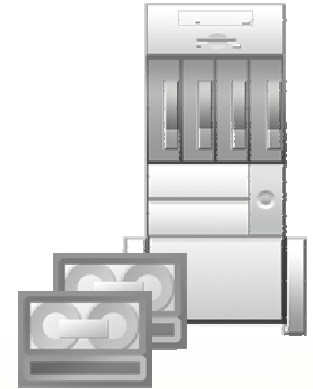


While tape remains the most common media type for storing compliant records, the capacity of records stored on disk-based systems will increase 172% between 2003 - 2006

The Archiving Challenge with Current Solutions

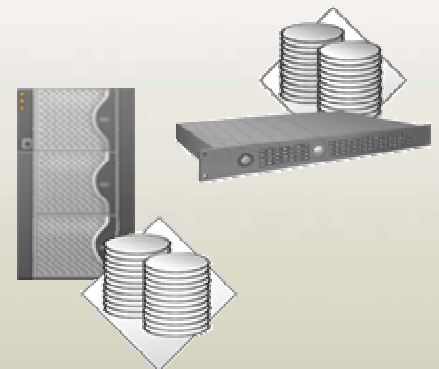
Archiving with Tape Systems

- ▶ **Data access is sequential – slower than online (disk)**
 - As the archive grows it becomes increasingly more administrative overhead
- ▶ **Archive data is tied to physical removable media**
 - Data is captive to that hardware system
- ▶ **Government regulations require long-term retention and authenticity**
 - Removable media is especially vulnerable to mishandling and corruption

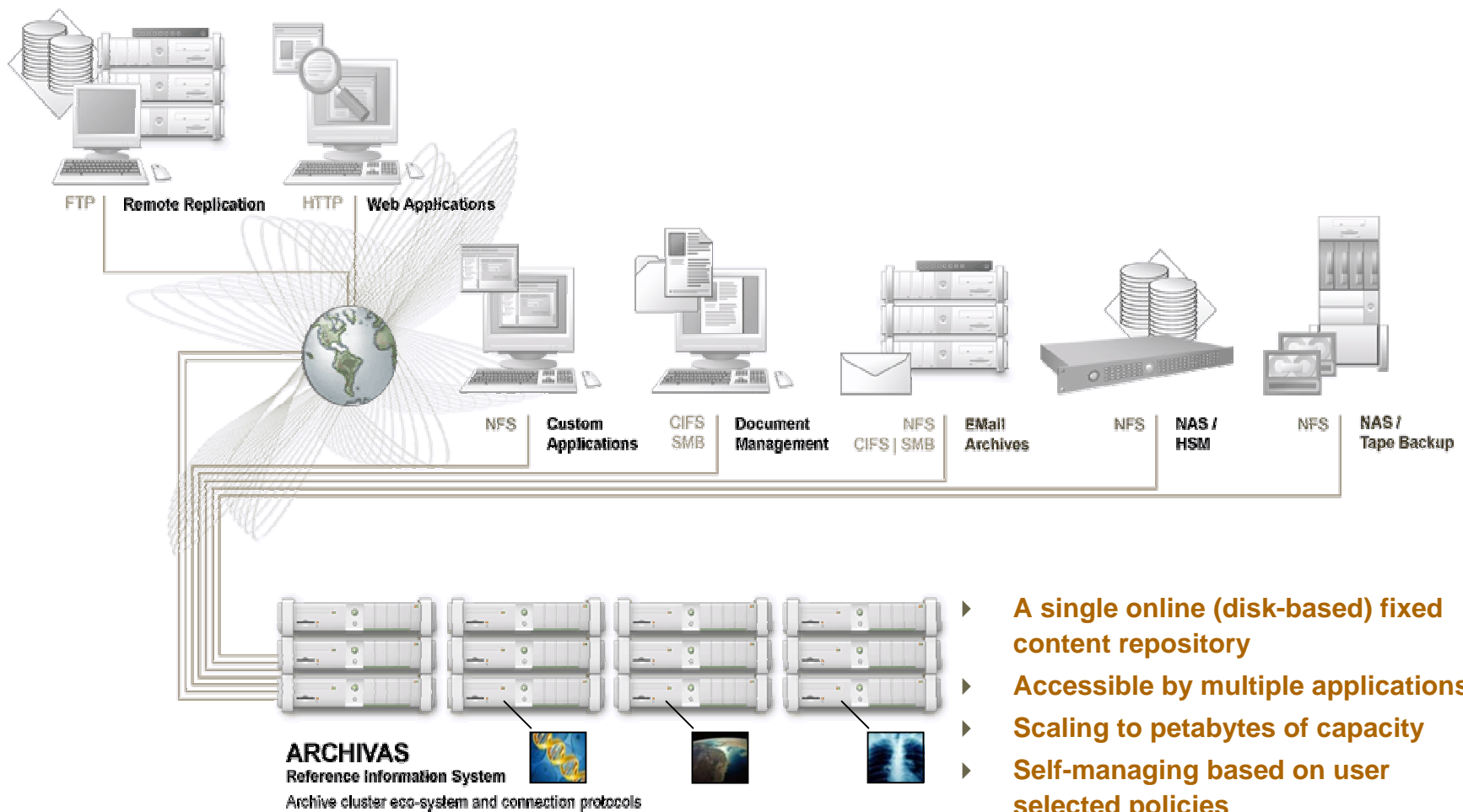


Archiving with Traditional RAID Systems or Network Attached Storage

- ▶ **Expensive to implement & manage**
- ▶ **Lacking in required WORM (Write-Once, Read-Many) and authenticity guarantees**
- ▶ **Optimized for high-performance but not ultra high (petabytes) scalability**



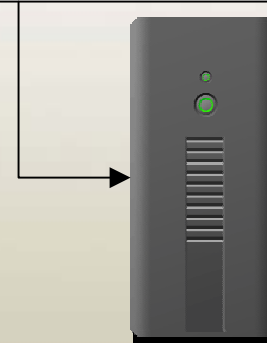
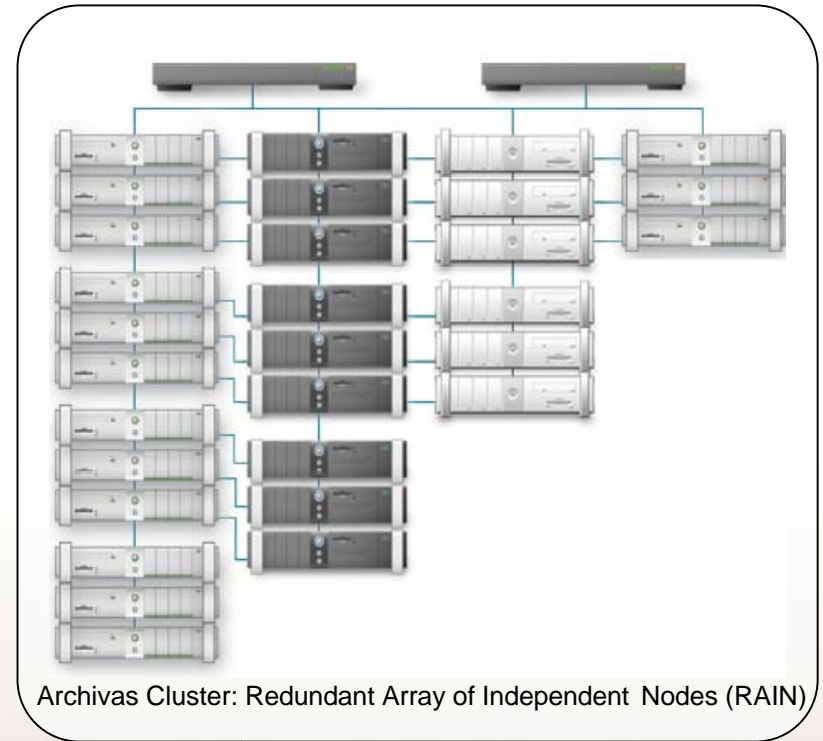
Requirement: An Online Archive



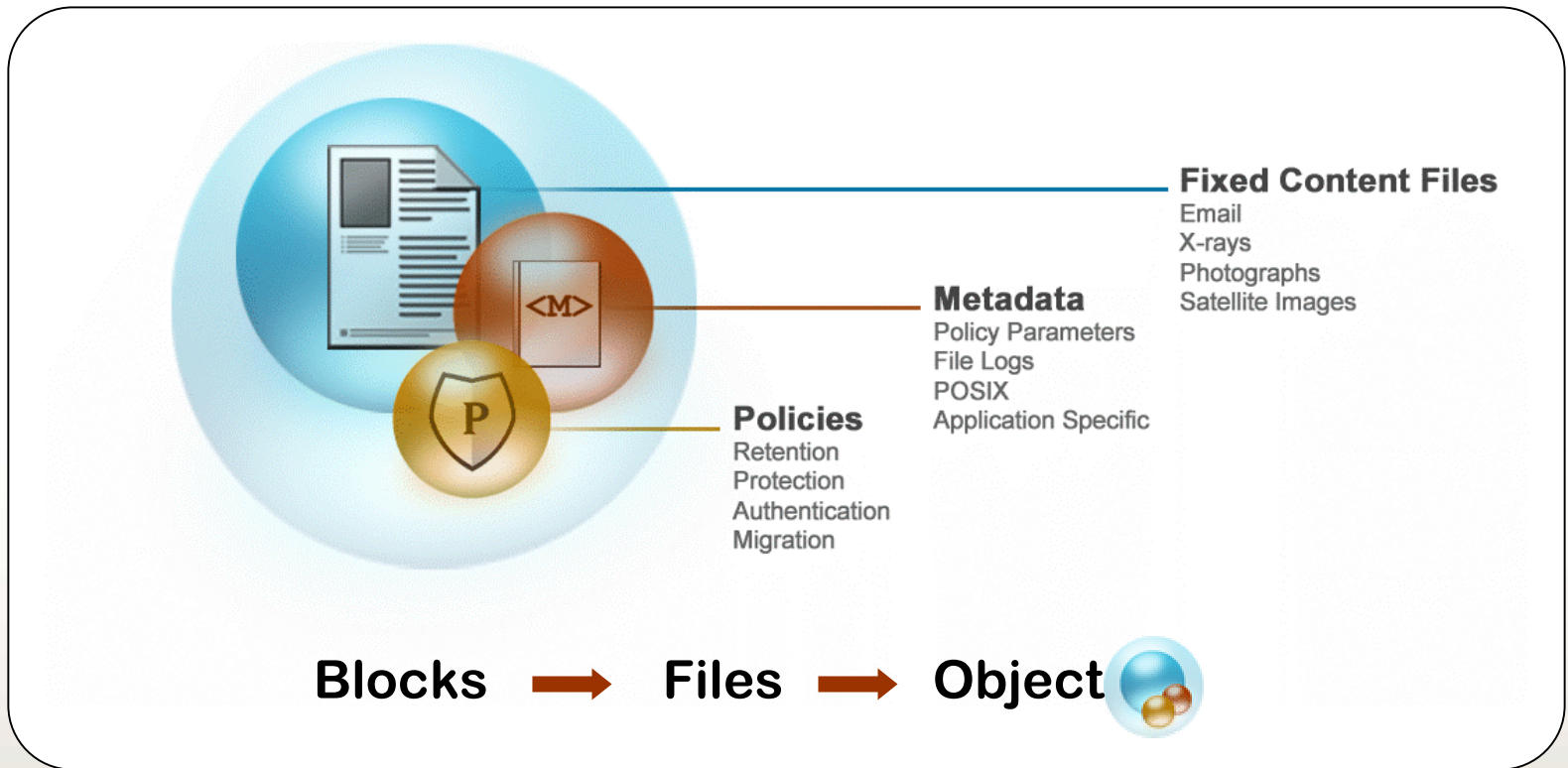
- ▶ A single online (disk-based) fixed content repository
- ▶ Accessible by multiple applications
- ▶ Scaling to petabytes of capacity
- ▶ Self-managing based on user selected policies

H-RAIN Architecture – Innovative Highlights

- ▶ **Object Based Storage**
 - Data
 - Metadata
 - Policies
- **Fully Symmetric Open Architecture**
 - Data store
 - Metadata management
 - Policy engine
 - Gateways
 - HTTP
 - NFS
 - CIFS / SMB
 - AFS
- ▶ **Support for heterogeneous hardware**
- ▶ **Self-configuring, self diagnosis and self healing, No Single Point of Failure (NSPOF)**



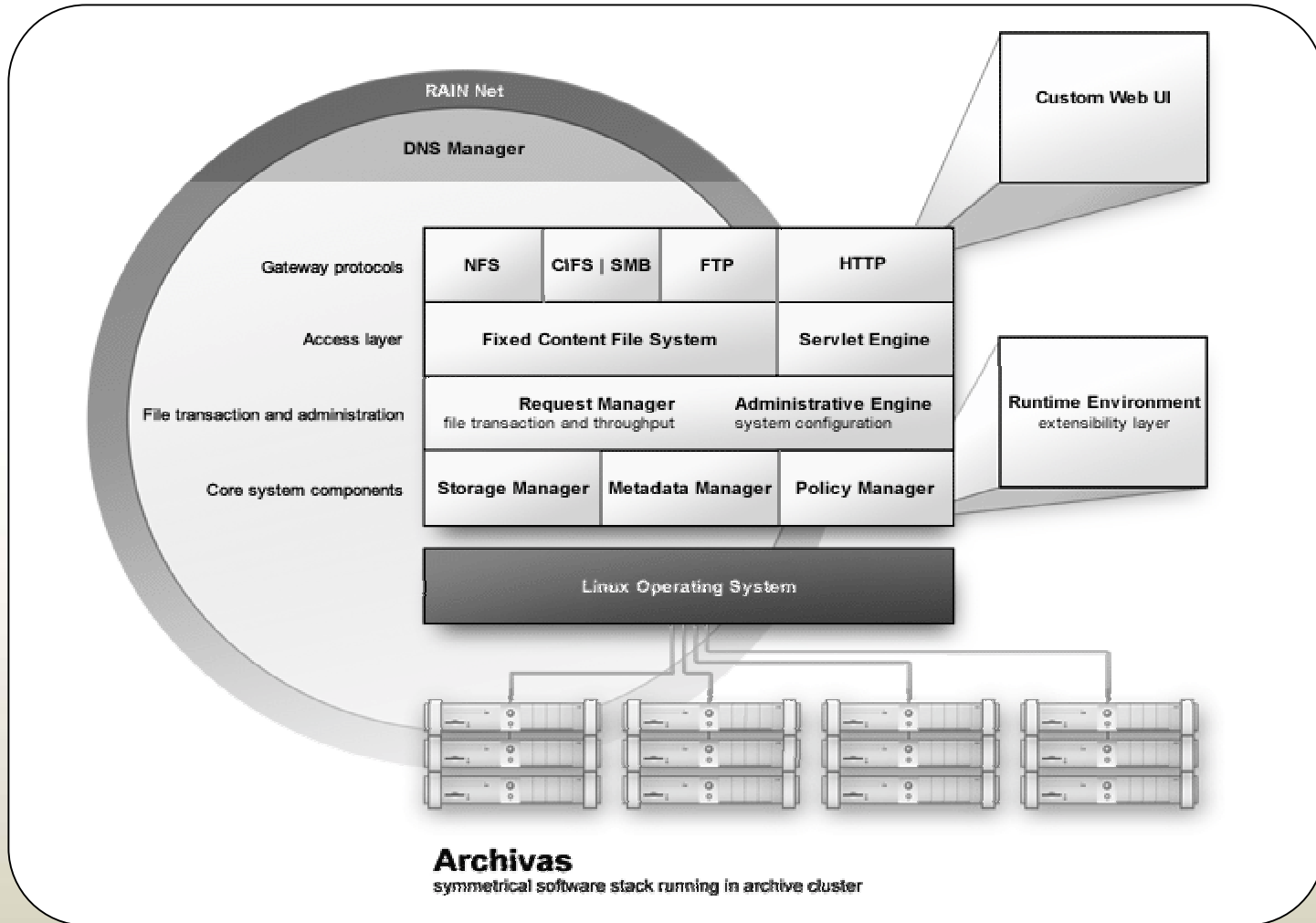
Creating & Managing “Archive Objects”



- ▶ **Grouping of data, metadata and policies** but not tied to specific applications
- ▶ Enable access and provisioning to multiple applications via standard gateways
- ▶ Increase data availability
- ▶ Insure data is backwards compatible and future proofed

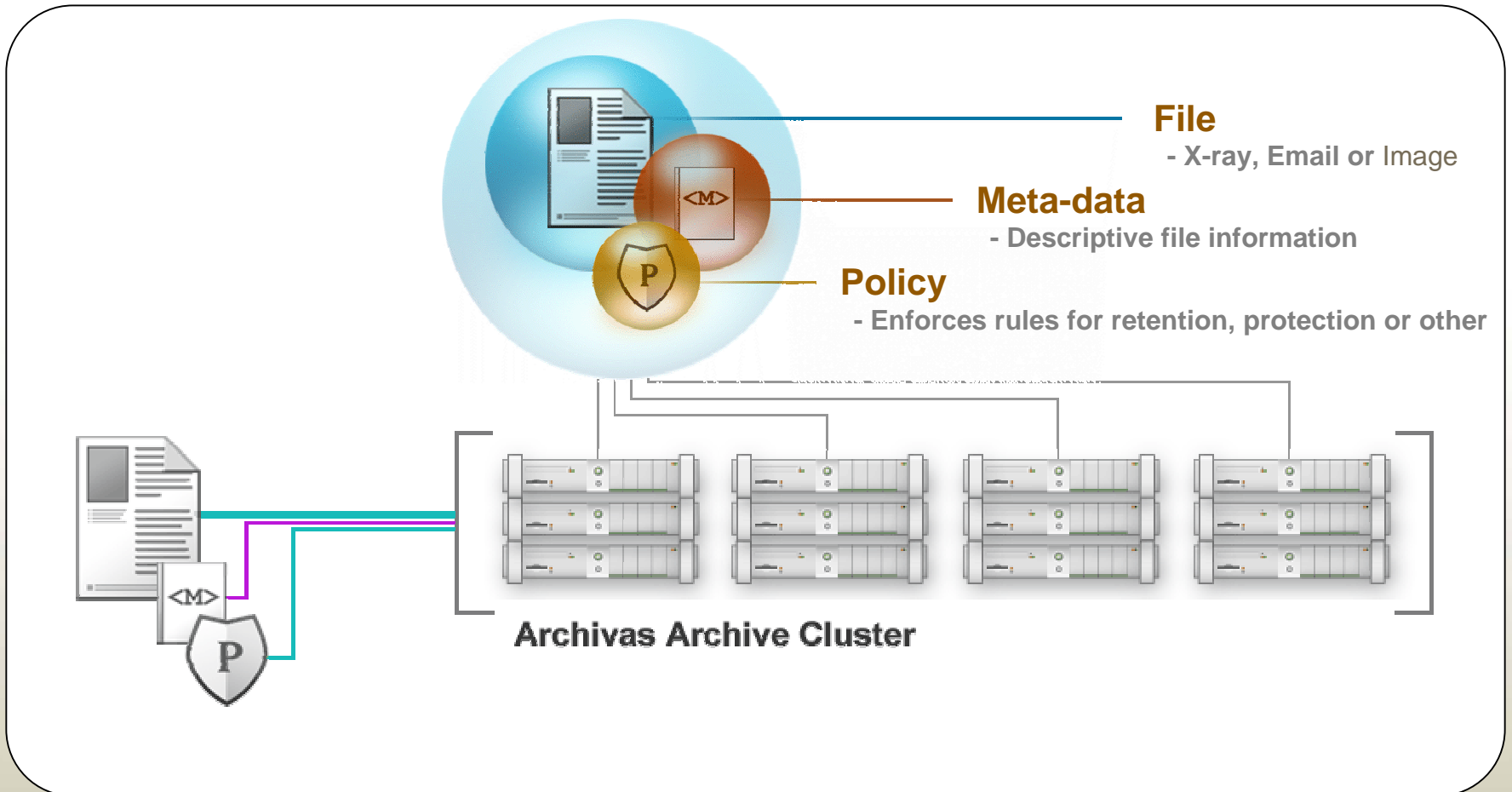
Archivas Architecture

Detailed view of an Archivas Cluster Node



Archivas Object-Based Management

- ▶ Keeps file, metadata and policies together
- ▶ Archivas stores and retrieves objects, not volumes or files



Archivas Fixed Content File System (FCFS)

- ▶ Management of archive objects
- ▶ 100% backwards compatible file system
- ▶ **Permanent association of data and metadata at the object level**

```
/legal/hr/offers/john_smith.doc
/legal/hr/offers/.../john_smith.doc/policy/
/legal/hr/offers/.../john_smith.doc/policy/authentication.xml
/legal/hr/offers/.../john_smith.doc/policy/migration.xml
/legal/hr/offers/.../john_smith.doc/policy/protection.xml
/legal/hr/offers/.../john_smith.doc/policy/retention.xml
...
/legal/hr/offers/.../john_smith.doc/history/file.xml
...
/legal/hr/offers/.../john_smith.doc/custom/john_smith.xml
```

Console
Screen

- ▶ Multiple metadata views (Symbolic links)
 - CAS
- ▶ Duplicate elimination (Hard links)
- ▶ Cluster monitoring and management

Archivas Object-Level Policies

▶ Retention

- Prevents deletion of a file before its retention period expires

▶ Authentication

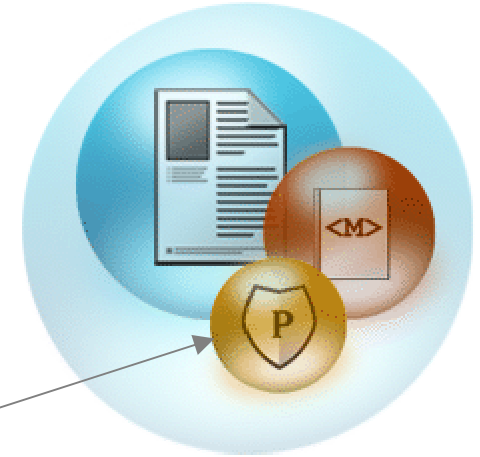
- Ensures that the content of a file matches its digital signature

▶ Data Shredding

- Overwrites file system entries where data objects were allocated with random fixed pattern data

▶ Protection

- Ensures the integrity of files, supports RAID-1, RAID-5 and RAID N+K



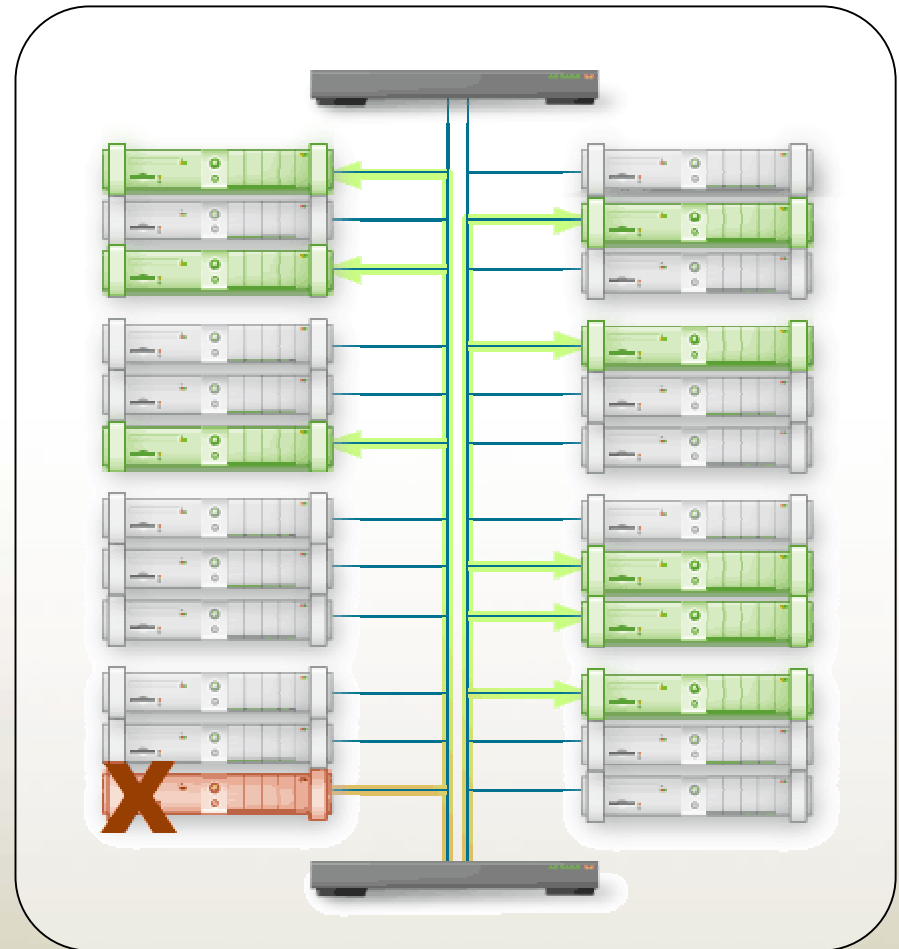
Archivas Cluster-Level Policies and Serviceability

Cluster policies:

- ▶ **Duplicate Elimination**
 - Insure no two copies of a single file are stored
- ▶ **Garbage Collection**
 - Reclaim disk space and perform integrity checks

Serviceability features:

- ▶ **Self-Protection**
 - Protection policies that enforce document retention, authentication, and file replication combine to protect an archive from loss of valuable digital assets
- ▶ **Self-Configuration**
 - Simplified installation and integration by setting systems configuration through high-level policies
- ▶ **Self-Balancing**
 - RIS reassesses and adjusts its own load balance by monitoring the activity and capacity of all nodes
- ▶ **Monitoring**
 - SNMP
 - FCFS

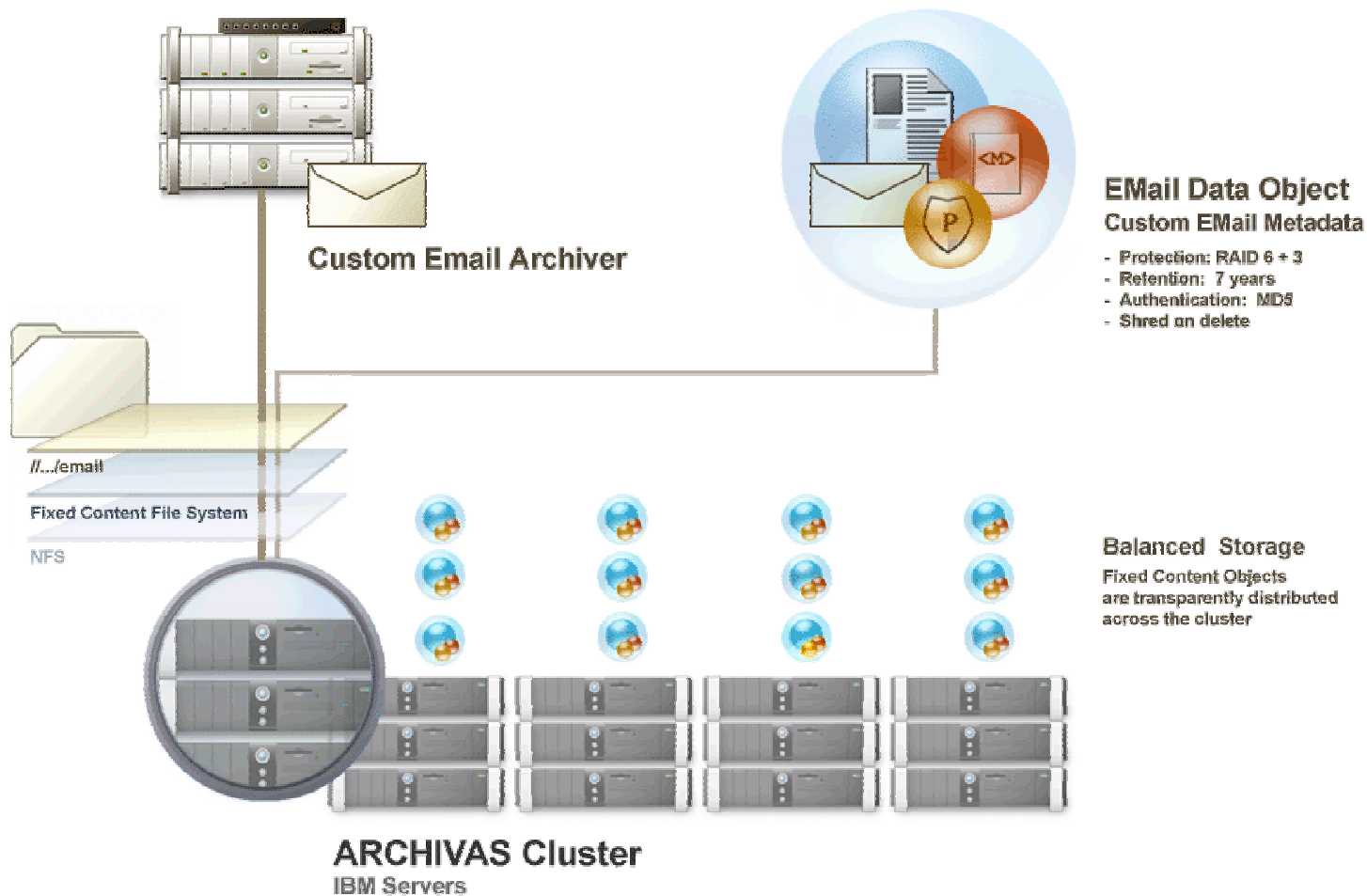


Use Case: Archivas in an enterprise environment

- ▶ Large amounts of fixed content
- ▶ Store data as unique archive objects
- ▶ Enable access to the archive by multiple applications
- ▶ Ability to use heterogeneous hardware/storage

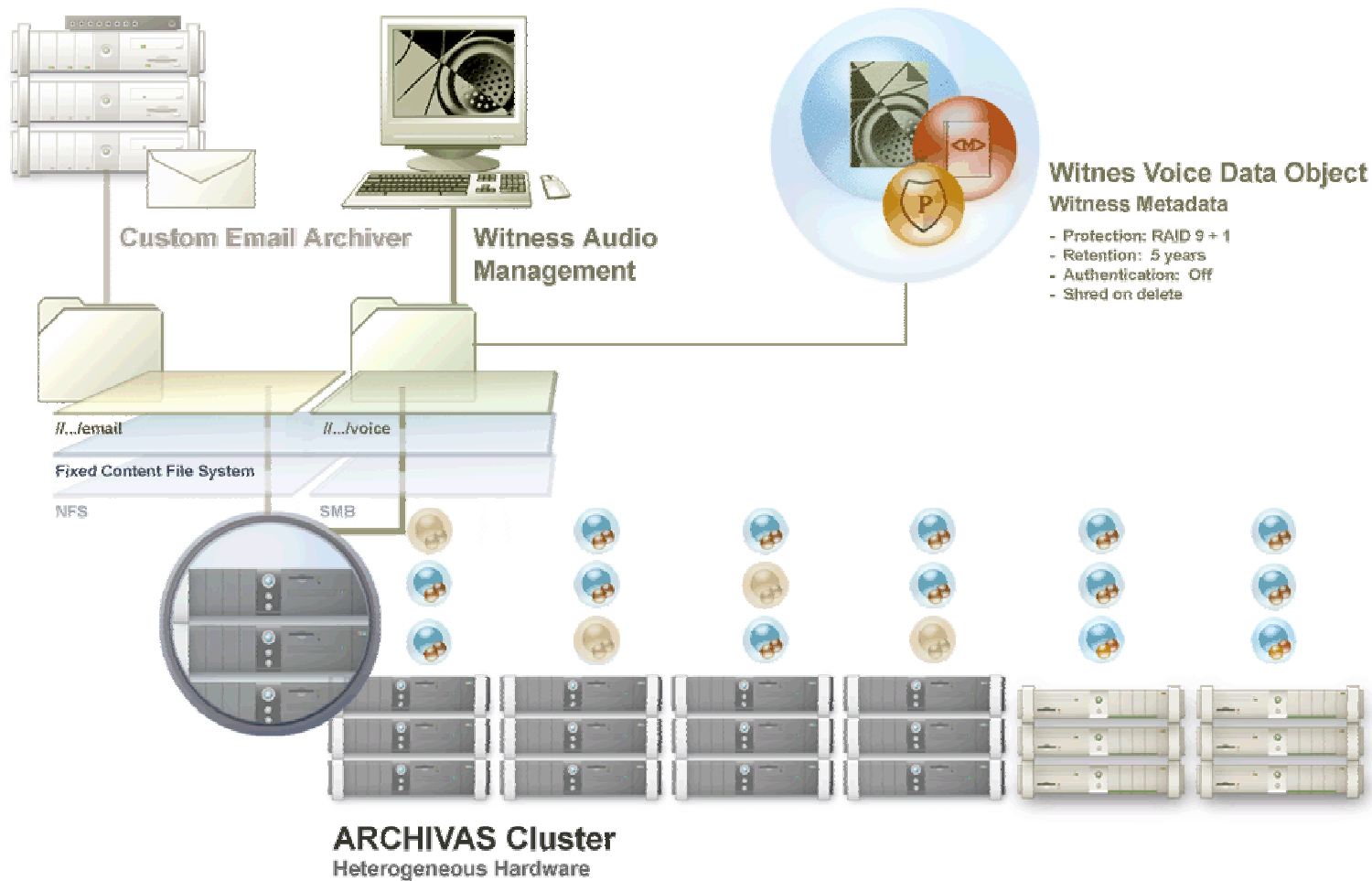
An Enterprise Use Case: Phase 1

- Single Application, Single Vendor Hardware Platform



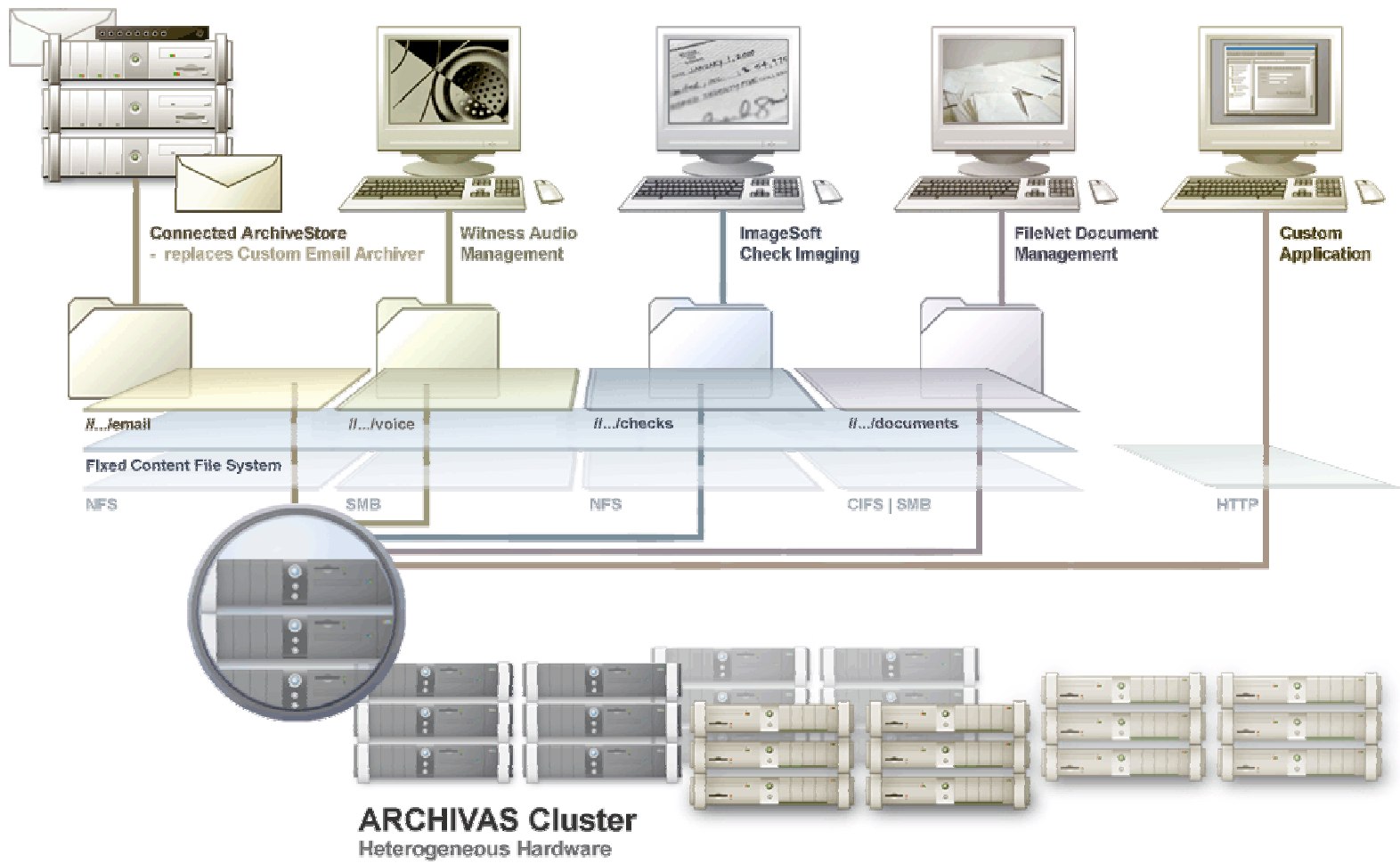
An Enterprise Use Case: Phase 2

- *Two Applications, Two Hardware Vendor Platforms*



An Enterprise Use Case: Phase 3

- Multiple Applications, Multiple Hardware Vendor Platforms



Summary: H-RAIN

- ▶ Object-based storage
- ▶ Fully symmetric open architecture
- ▶ Support for heterogeneous applications and hardware
- ▶ Self-configuring, self diagnosis and self healing, No Single Point of Failure (NSPOF)