Electronic Records Management and Retention

"…The challenge is to determine which types of electronic records must be retained and for how long, as well as how to best preserve them and make them available…"

Presenters

David Warner
Sr. Solutions Specialist – The Americas
Records Management Products
Content Management & Archiving (CMA) Division
EMC

Tim Baker
Deputy State Archivist
Maryland State Archives

Role of the Archives

Historical agency for Maryland serving as central repository for permanent records - Records date from the founding of the colony in 1634 and include:

- Executive, legislative and judicial records
- County and Municipal
- Probate
- Land records
- Court records
- Church records

- Business records
- Publications and reports
- Special collections:
  - Maps
  - Photos
  - Private papers
  - newspapers

My background...

- 5 years as Deputy at MSA

- 2 Years at Enterprise Information Solutions

- 6 years at Maryland Department of Planning

- 9 years in the Governor's Office

- 2 years in Office of the Mayor of Baltimore

- 4 years on Capital Hill

What is a Record?

"public record"

- "original or any copy of any documentary material ... made ... or received by [an agency] in connection with the transaction of public business" SG §10-611(g)(1)(i)

- can be in "any form", including without limitation: card, **computerized record**, correspondence, drawing, film or microfilm, form, map, photograph or photostat, recording, tape SG §10-611(g)(1)(ii)

- "public record" includes both printed and electronically stored versions of e-mail messages, e-mail messages never printed out, and includes e-mail messages related to agency business on employee's home computer.

81 *Opinions of the Attorney General* 140 (1996)

Where is it defined in statute?

Maryland Public Information Act ("MPIA")

Annotated Code of Maryland, State Government Article ("SG"), §10-611 *et seq.*

Reference: *Public Information Act Manual* (10th Edition 2007)

www.oag.state.md.us/opengov/index.htm

Why have a records management program?

- Necessary for effective governance
- Protects the rights of citizens and interests of the agencies
- Accountability
- Financial records document how tax dollars are spent
- Directives and reports document what you do and how well you do it
- Notes, minutes and memos document how decisions are made
- You might be fined, fired or both !!!

Problems Records Management

- Agencies tend to think of records in terms of paper files or series of paper files

- Therefore, record retention programs tend not to include vitally important electronic record series

- Agencies have tried to hold onto records they designate as permanent

Record series not considered "records"

- Access Log Files
- Security Log Files
- Voice Mail
- email
- databases

*** A record is ANYTHING created by or received by an agency ***

What do you need to do?

- Make sure your agency has a viable records management policy and procedure

- Educate your IT staff on basic records management requirements

- Inventory all information assets under control of the IT operation

- Identify agency records manager or appoint an IT records manager

- Incorporate the inventory into the records retention and disposition schedule

Electronic Records Management

Goals:

- integrate e-records management concepts and practices with comprehensive information management policies, processes and objectives to assure the integrity of e-records and information.

- employ ERM to support interoperability, timely and effective decision making, and improved services to customers.

- provide the tools for agencies to access e-records for as long as required and to transfer permanent e-records to MSA for preservation and future use by government and citizens

Electronic Records Management

- State Wide Initiatives We Need to Work On...

Development of standards for preservation of the permanent record (pdf – a)

Design and architect means to migrate permanent record material to the archives immediately upon creation (xml may be the way)

Electronic Records Management

- Initiatives

Interoperability   **Continuity**   **Public Access**
Preservation

MSA is working with many agencies to establish some model projects for interoperability, continuity, public access and preservation

Partnerships with
         UMBC for remote site
         Judiciary for land records
         SHA Right of Way Plats
         Montgomery County Road plans
         Prince George's County (MNCPPC) GIS / plats integration
         SDAT linkage from land records to assessment files and visa versa

We all have an obligation to work together to document the work we do, ensure a high level of accuracy and availability of information and data critical to our operations and to preserve the people's records.

Some issues with data...

- Collection and transfer
  - Open standards for transfer and retention ???
  - Proprietary applications which lack export function

- Conversion, assurance, consolidation and integration

- Integrity - - monitoring and audit data flow

- Security

more issues with data...

- Correction and expungement

- Delivery and sharing
  - As well as sheilding and restricting

- Certification

The Model: mdlandrec.net

A joint initiative of the Judiciary and the Archives

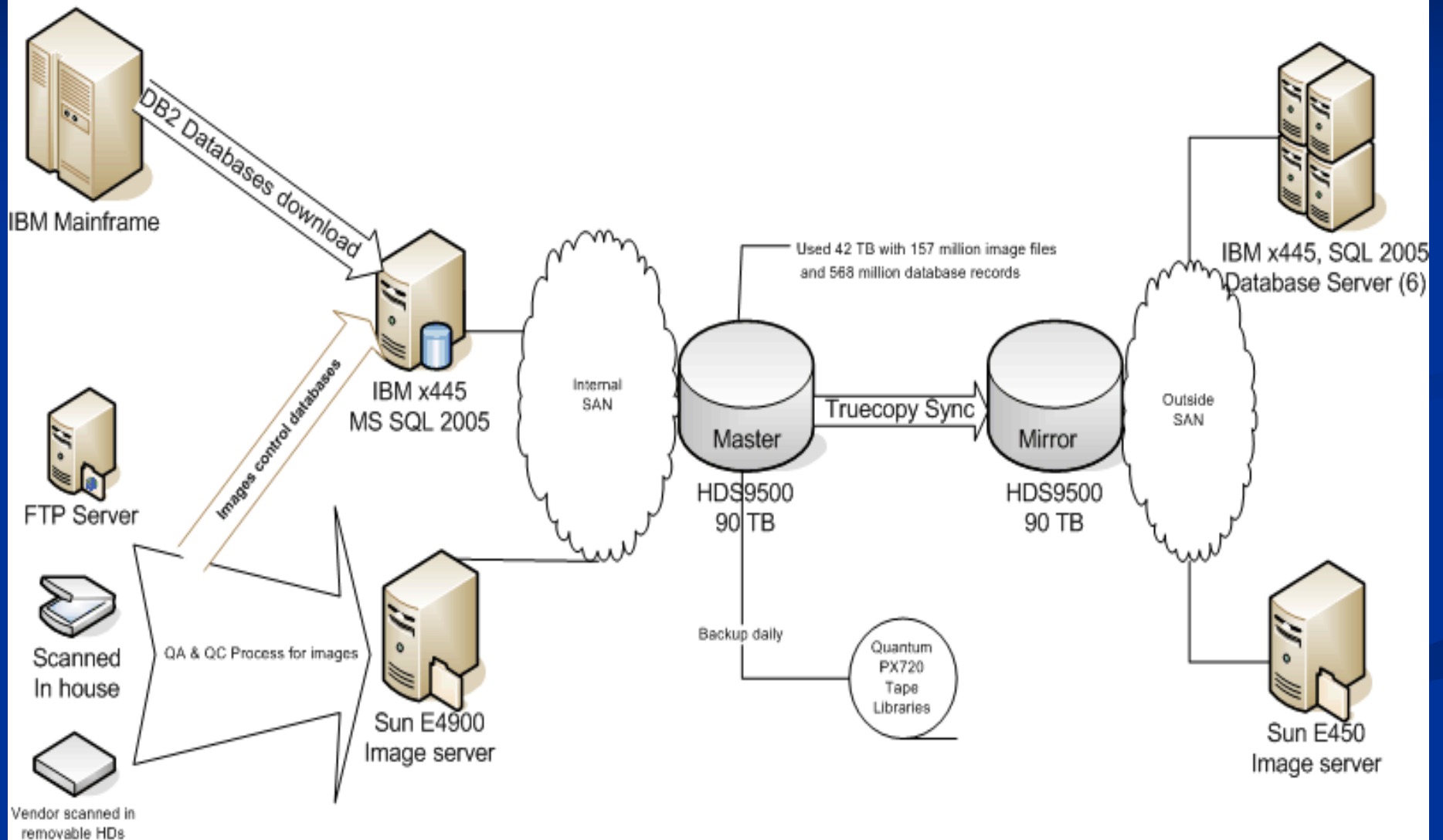... "a model system accessing and preserving permanent electronic records."

Components: 170,000,000 images

500,000,000 database records

www.mdlandrec.net

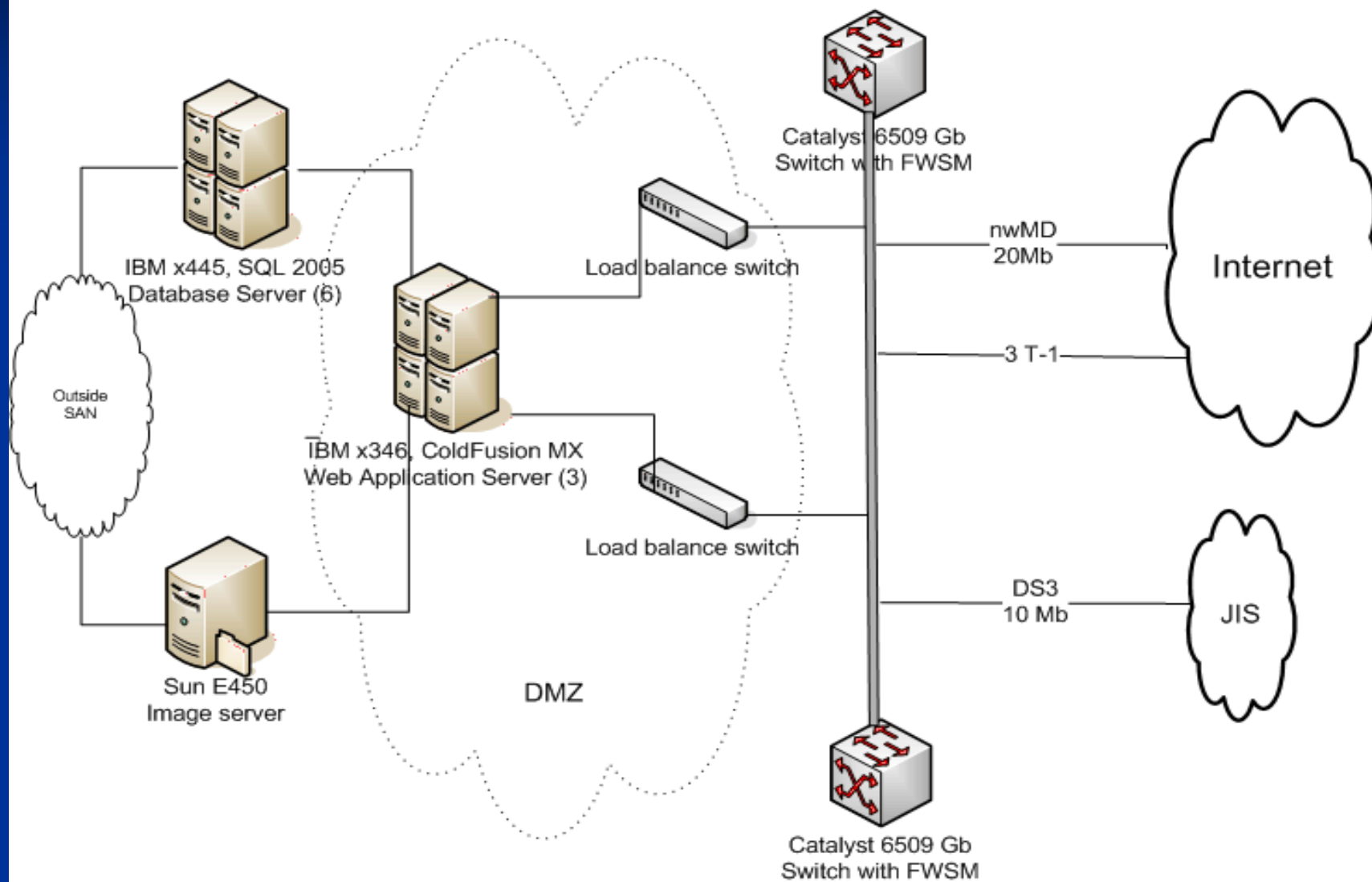# Maryland State Archives ---- MDLandRec.net Project
Data Collection and Internal Process

IBM Mainframe

DB2 Databases download

IBM x445
MS SQL 2005

FTP Server

Images control databases

Scanned
In house

QA & QC Process for images

Vendor scanned in
removable HDs

Sun E4900
Image server

Internal
SAN

Used 42 TB with 157 million image files
and 568 million database records

Master

HDS9500
90 TB

Truecopy Sync

Mirror

HDS9500
90 TB

Backup daily

Quantum
PX720
Tape
Libraries

Outside
SAN

IBM x445, SQL 2005
Database Server (6)

Sun E450
Image server

Maryland State Archives ---- MDLandRec.net Project
Data Retrieval and Delivery

Data Synchronization

Agency Database

Record Number	Modification Date
1	5-01-2007
2	5-17-2007
3	5-09-2007
4	5-31-2007

Archives Database

Record Number	Modification Date
1	5-01-2007
2	5-04-2007
3	5-09-2007

Synchronizing two separate databases is achievable if both databases follow a few rules. Each table that is to be synchronized needs to have two fields, Record Number and Modification Date.

Record Number is a unique number that is a permanent part of that record. This field is not allowed to change or be unknown. The uniqueness of this field only exists throughout this table.

Modification Date represents the last time any of the data in that record has been changed. It is required that this field is updated with the current date and time whenever the record is altered.

Steps in a data synchronization application executing on an Archive's server:

1. Determine the Archive's Last Modification Date, such as 5-9-2007.

2. Request only the records from the Agency Database that are later than the Archive's Modification Date.

3. The Agency Database sends the records that have a Modification Date later than 5-9-2007 to Archives. In this case, Record Numbers 2 and 4 would be transferred.

4. In the Archives Database, Record Number 2 would be replaced and Record Number 4 would be added, bringing the database into Synchronization.

5. Repeat for all tables within the Agency Database.

Data Verification

Agency Database

Record Number	Modification Date
1	5-01-2007
3	5-09-2007
4	5-31-2007

Archives Database

Record Number	Modification Date
1	5-01-2007
2	5-17-2007
3	5-09-2007
4	5-31-2007
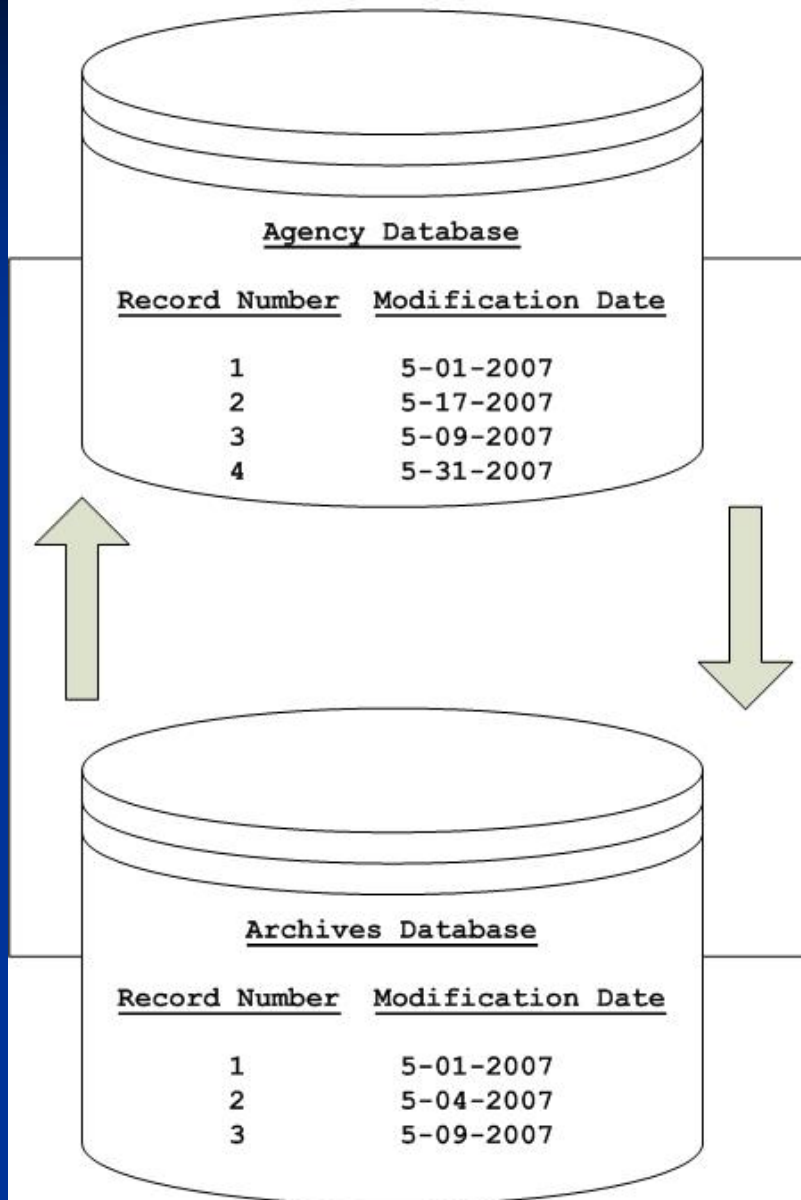
After synchronizing two databases, it will be required to verify that those databases have been correctly synchronized. To achieve this, there are two main types, Fast Verification and Full Verification.

Fast Verification requires a count of the total number of records in each of the databases. If the two counts match, it is assumed that the databases are identical. If the Agency Database has more records, then an algorithm is used to find the missing Record Numbers. Any missing records are then transferred and added to the Archives Database. If the Agency Database has less records than the Archives Database, then a similar algorithm is used to find the missing Record Numbers, which are promptly removed from the Archives Database.

Full Verification requires a complete download of the Agency Database which is then compared to the Archives Database on a Record by Record basis. Any changes that are needed to bring the Archives Database into synchronization are then performed. This verification is clearly the best, but is also the most resource intensive type as well.
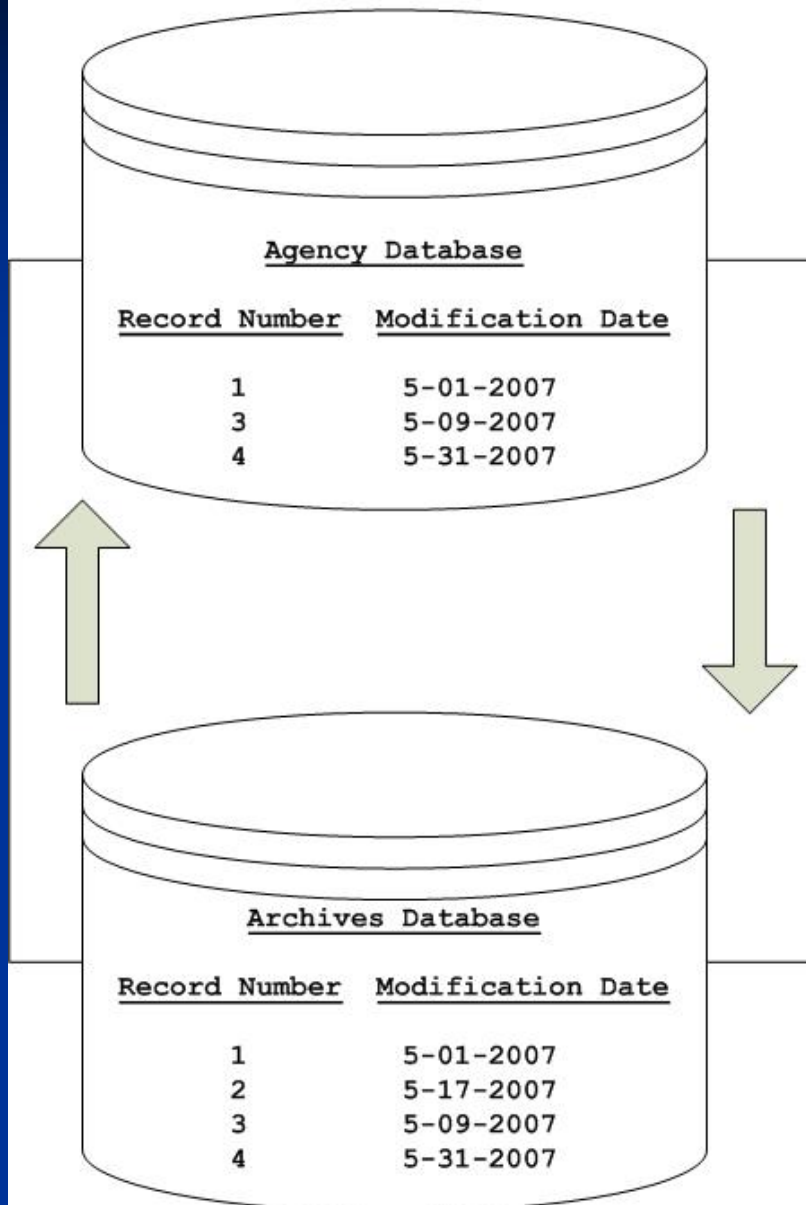
Data Replication as a Preservation Tool

- Application or Host based replication
- Array Based Replication
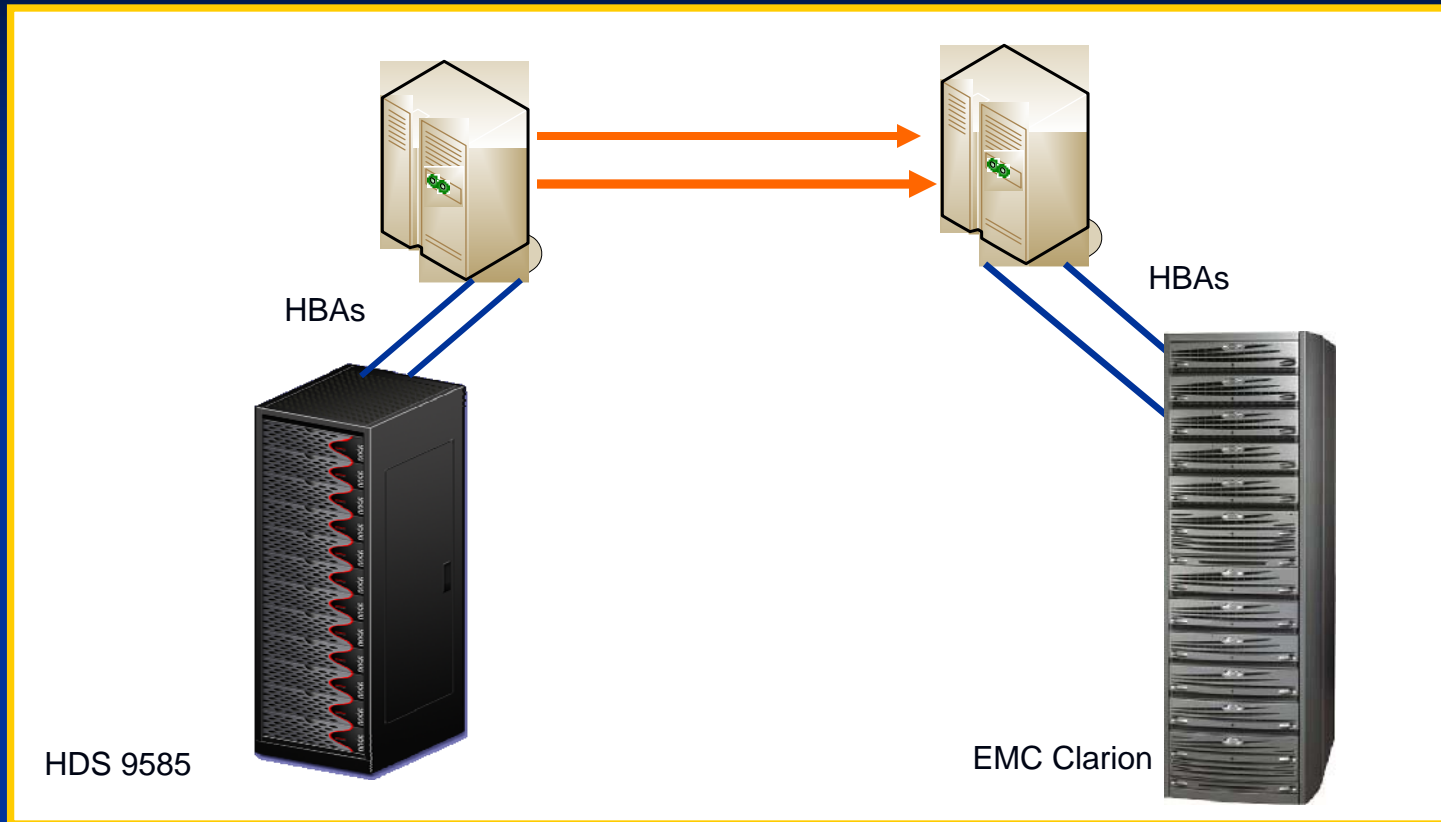- Appliance based replication

Array Based Replication

SunFire
4900

SAN

SAN

Sun E450
Image server

HDS TrueCopy – 2GB HBAs
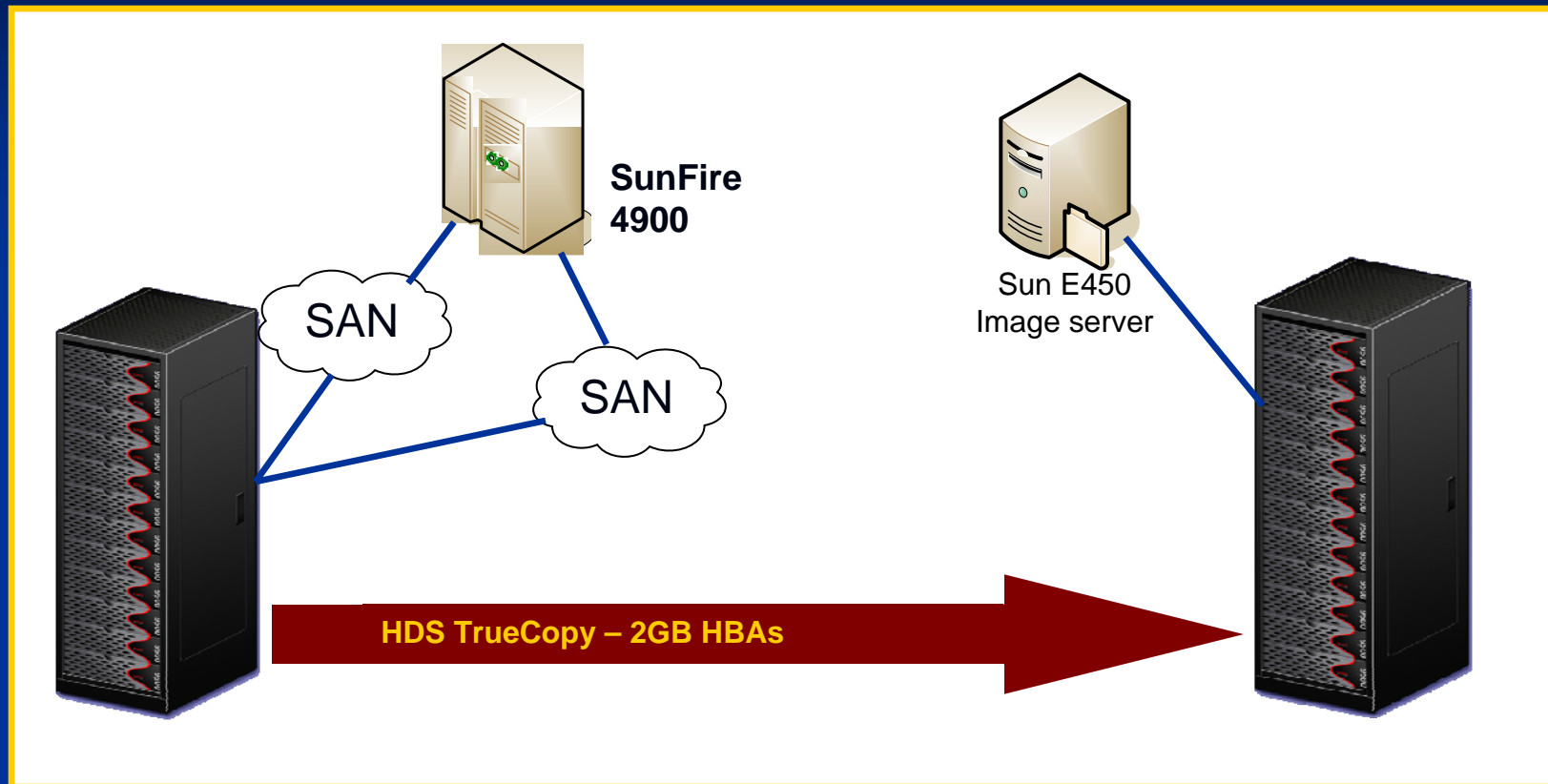
Replication takes place at the controller level - - bit by bit

Must be homogeneous

Distance Issues

Host Based Replication

HBAs

HBAs

HDS 9585

EMC Clarion
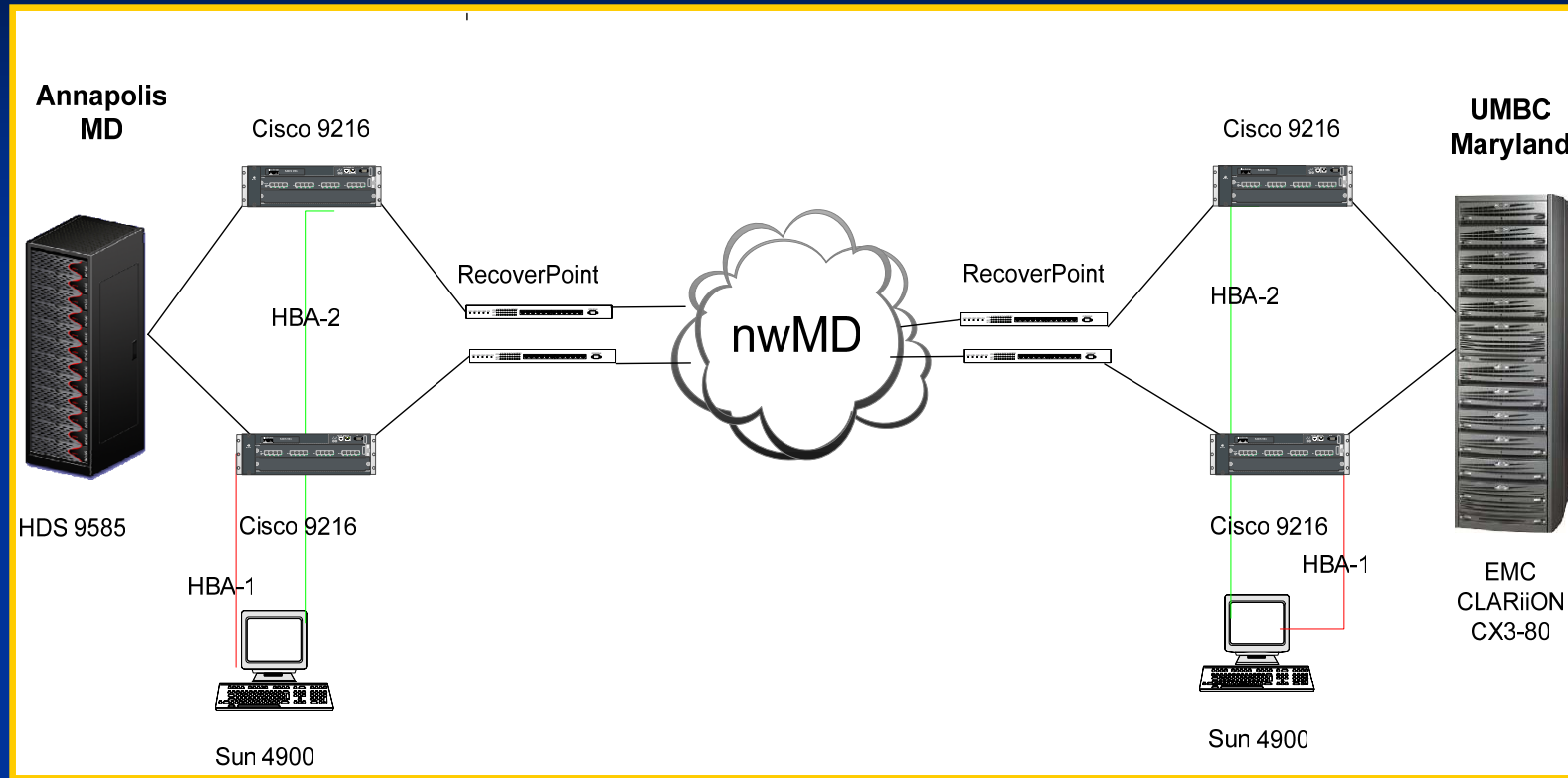
Allows heterogeneous storage

Open standard but all writes go through OS and hosts

Appliance based replication

Solution offers dedicated host

Open standard

Wrap up

- Know what records you have

- Know when and how you can dispose of non-permanent records

- Develop a record retention schedule and identify permanent records

- Work with the Archives to ensure safe transfer preferably at time of creation