



NATIONAL
ARCHIVES
OF AUSTRALIA

Access Across Time: How the NAA Preserves Digital Records

Andrew Wilson
Assistant Director, Preservation



NATIONAL
ARCHIVES
OF AUSTRALIA

What I will talk about

- NAA Context
- Some Concepts
- NAA Implementation
- NAA Process flow
- Preservation Software Platform (Xena)



NATIONAL
ARCHIVES
OF AUSTRALIA

National Archives of Australia

- Established 1946 as part of National Library
- Independent since 1960
- Legislation: Archives Act 1983
- Approx. 420 staff in 12 locations
- Budget ca. A\$65 million.
- 350 shelf kilometres of records
- Separate preservation funding since 2001



NATIONAL
ARCHIVES
OF AUSTRALIA

Digital Preservation Project

- Started in 2001/2 FY
- Cost to date approx. A\$2 million (30 months)
- Aim: to develop viable approach to the preservation of 'born digital' records for long term accessibility and use
- Deadline: July 2003



NATIONAL
ARCHIVES
OF AUSTRALIA

Note

- NOT a digital archive but an approach to digital preservation
- Well developed archival processes that can be applied to records irrespective of format:
 - appraisal/selection
 - transfer
 - description
 - retrieval and access
- Project purely about preservation



NATIONAL
ARCHIVES
OF AUSTRALIA

Some Definitions

- Records

Recorded information created or received and maintained by an organisation in the transaction of business

- Digital Records

Records in digital form processed by computers

- *Not:*

Systems or working applications



NATIONAL
ARCHIVES
OF AUSTRALIA

The preservation problem

- Technological obsolescence
 - Hardware
 - Software
- Restrictions on the use of technology



NATIONAL
ARCHIVES
OF AUSTRALIA

Traditionally

Researcher directly experiences the record through its source object



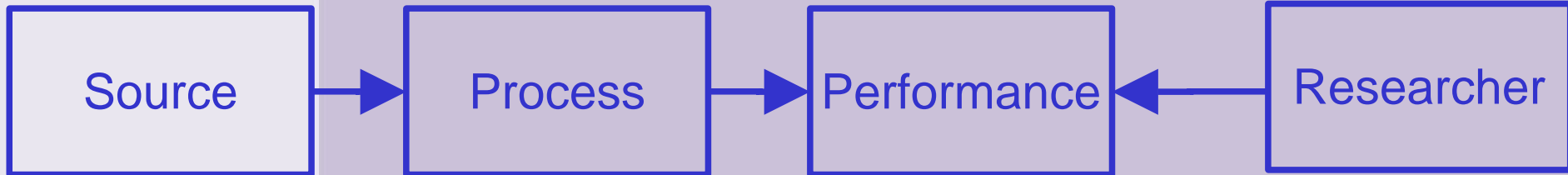
Preserve the object and you preserve the record



NATIONAL
ARCHIVES
OF AUSTRALIA

But...digital records are performances

Researcher experiences the record
through a performance



*Preserve the performance and you
preserve the record*



NATIONAL
ARCHIVES
OF AUSTRALIA

A Two Part Solution

1. Keep a *master copy* of every source we accept into custody
 - *Passive Access*
 - Researcher gets the 'Zeros and Ones', not the performance
2. *Active Intervention* to recreate the performance
 - Replace the source and process
 - *Active Access* to the 'essence' of the performance
 - Based on experience with Audiovisual material



NATIONAL
ARCHIVES
OF AUSTRALIA

The *essence* of the record

- What we want to preserve out of the performance
 - What aspects are essential to the record's value?
 - What aspects are incidental to the record's value?



NATIONAL
ARCHIVES
OF AUSTRALIA

Our preservation approach

- Select open and well documented data formats
- Migrate records into these formats ('normalisation')
- Support open source software tools that can read these formats



NATIONAL
ARCHIVES
OF AUSTRALIA

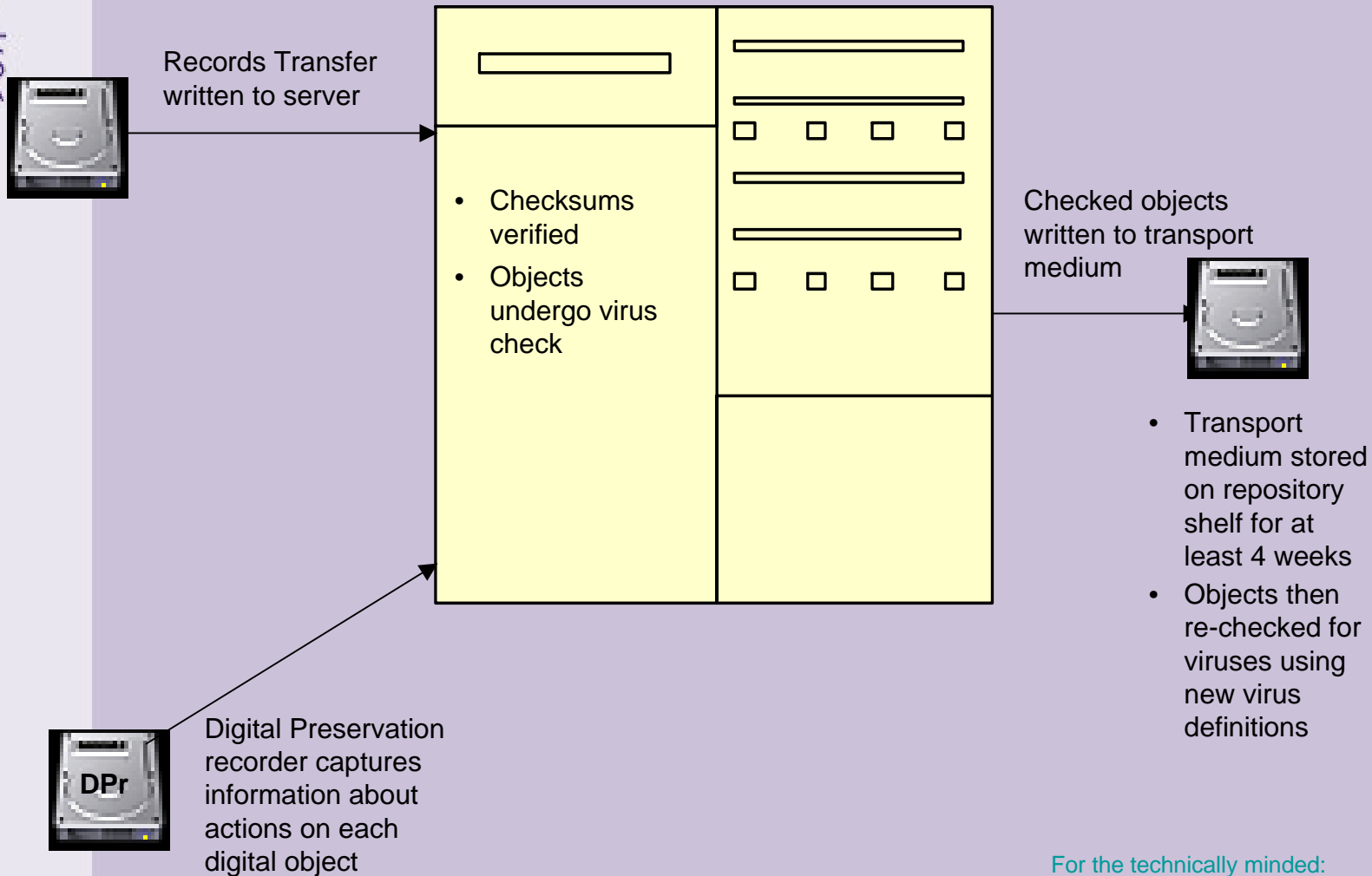
Preservation System

- 3 separate components
 1. Quarantine
 2. Preservation
 3. Storage
- All components physically separated from each other and all other NAA networks
- Access to hardware restricted to digital preservation staff



NATIONAL
ARCHIVES
OF AUSTRALIA

Quarantine server

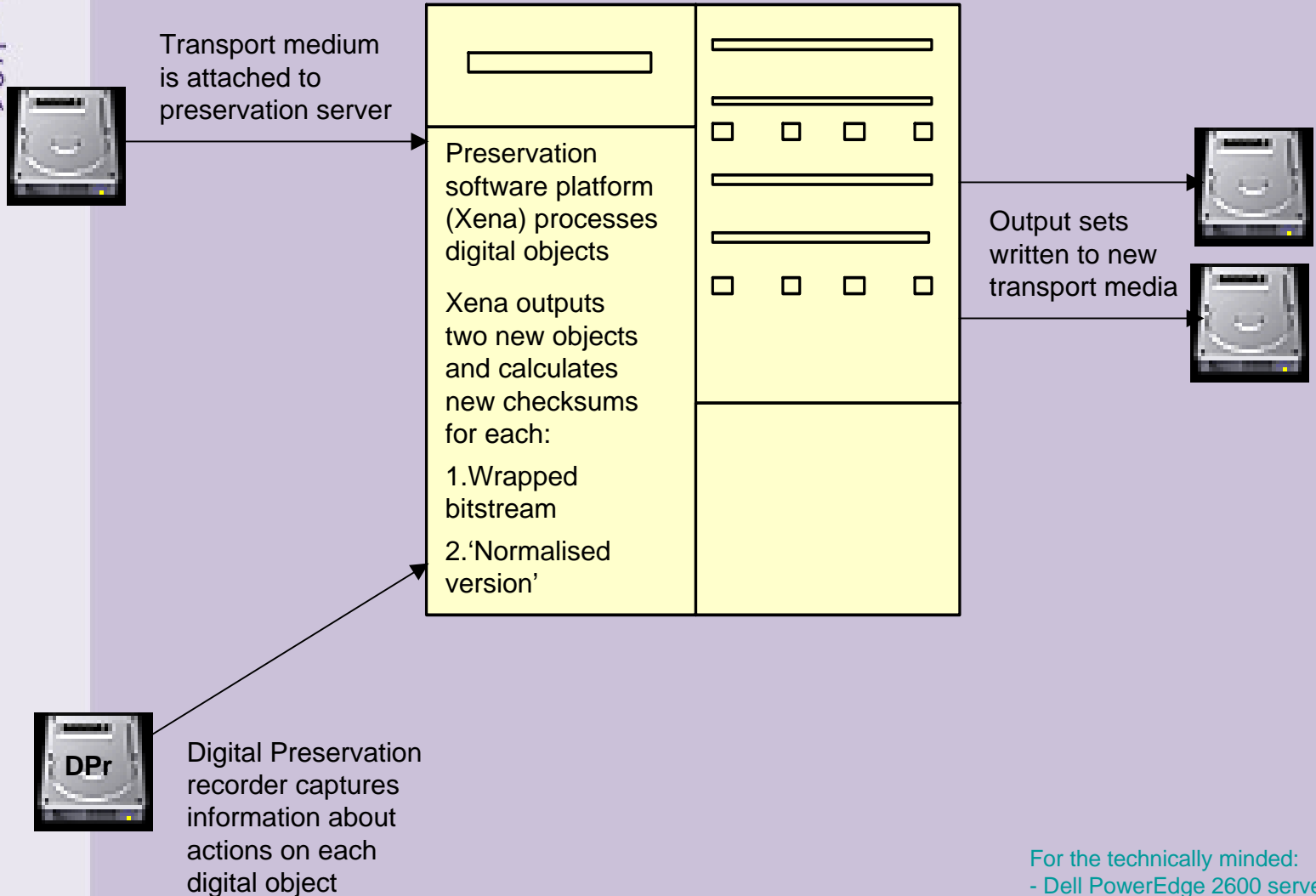


For the technically minded:
- Dell PowerEdge 2600 server
- 2 x 2GHz processors
- .7Tb disk store
- independent UPS



NATIONAL
ARCHIVES
OF AUSTRALIA

Preservation server



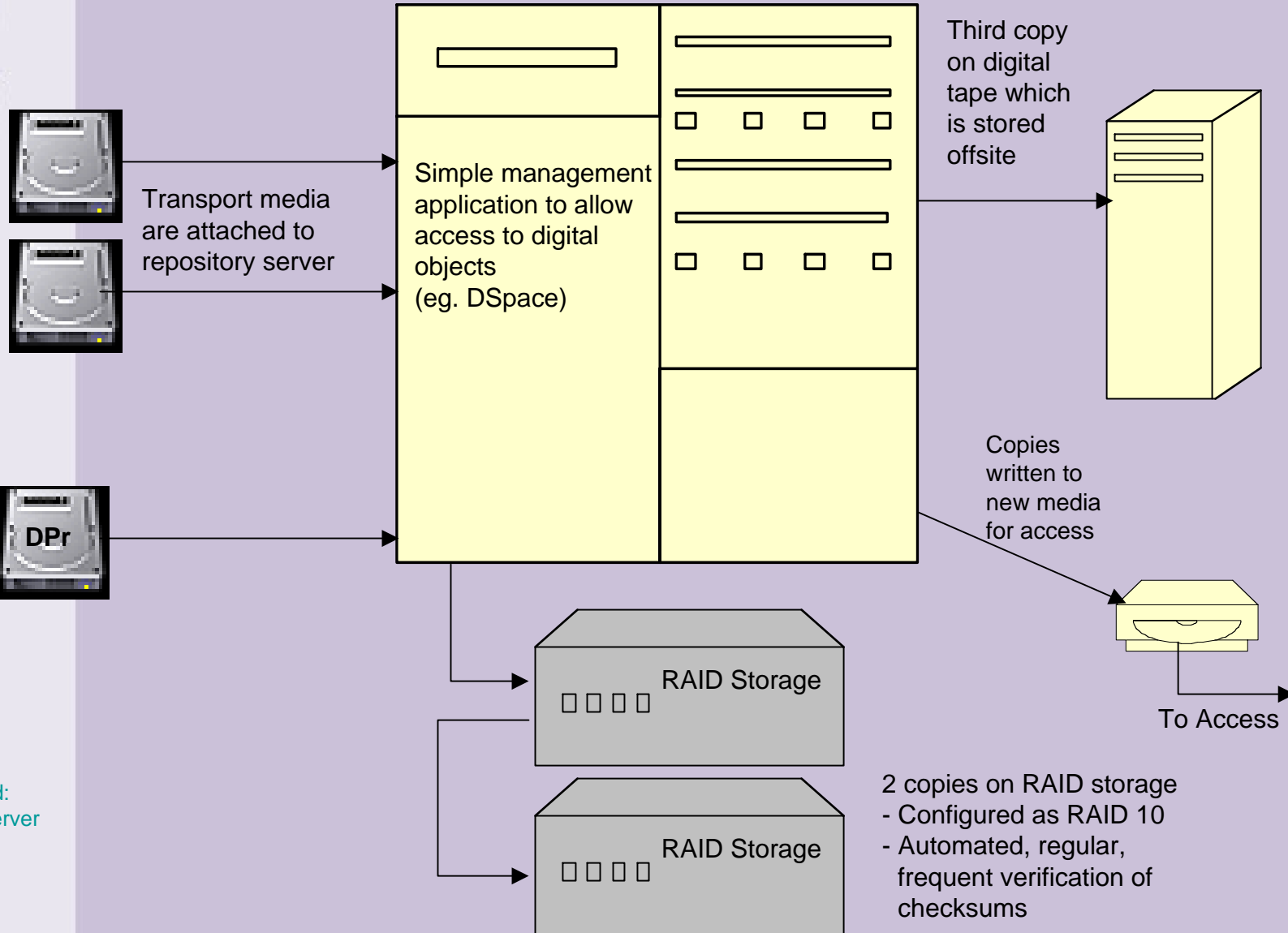
For the technically minded:

- Dell PowerEdge 2600 server
- 2 x 2GHz processors
- .7Tb disk store
- independent UPS



NATIONAL
ARCHIVES
OF AUSTRALIA

Digital Repository



Digital Preservation recorder captures information about actions on each digital object

- For the technically minded:
- Dell PowerEdge 2600 server
 - 2 x 2GHz processors
 - .7Tb disk store
 - fibre channel between server and RAID
 - independent UPS



NATIONAL
ARCHIVES
OF AUSTRALIA

NAA Implementation

1. Follows Open Archival Information System framework
2. Non-proprietary, open source solution
3. Based on the extensible markup language (**xml**)



NATIONAL
ARCHIVES
OF AUSTRALIA

xena

=

xml electronic normalising
of archives





NATIONAL
ARCHIVES
OF AUSTRALIA

xena

- File-based
- Java/Swing application
- Runs in Java 1.3 +
- Packaged as an executable .jar file
- Modular
- Multiple document interface



NATIONAL
ARCHIVES
OF AUSTRALIA

xena functionality:

A core module plus 'plug-in' modules which do:

- File format guessing
- File 'normalisation'
- XML encapsulation
- Process and data verification
- File viewing



NATIONAL
ARCHIVES
OF AUSTRALIA

Core module

The core consists of:

- I. Graphical User Interface components
- II. Plug-in management components
- III. Generic validation components



NATIONAL
ARCHIVES
OF AUSTRALIA

Plug-in modules

- Plug-ins are created for identified data types that are to be processed. Each plug-in consists of:
 - A guesser component
 - One or more input format type components
 - A normalised format type
 - One or more normalisation modules
 - One or more view components
 - Sorting functionality
 - Validation functionality
 - Printing functionality
 - GUI interaction methods



NATIONAL
ARCHIVES
OF AUSTRALIA

DEMONSTRATION OF XENA



NATIONAL
ARCHIVES
OF AUSTRALIA

Contacts

Andrew Wilson
Project Manager
AtoR Digital Preservation Project
+61 2 6212 3694
andreww@naa.gov.au

Web:
<http://www.naa.gov.au/recordkeeping/preservation/summary.html>



NATIONAL
ARCHIVES
OF AUSTRALIA

THANK YOU

ANY QUESTIONS?

