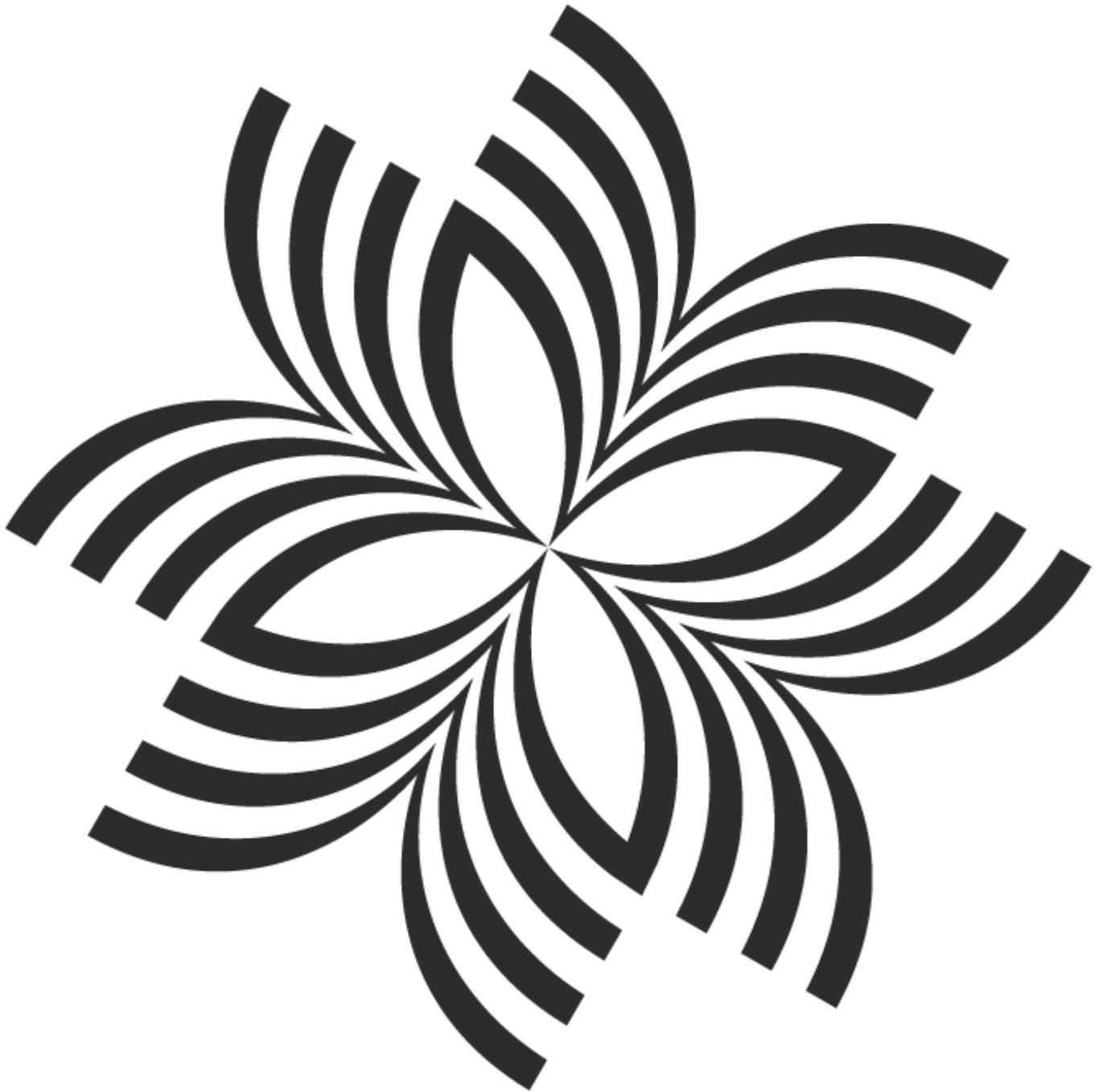


# National Digital Heritage Archive Programme Appendices

Appendices to the NDHA Business Requirements Specifications



## Document control

### Revision history

<b>Revision</b>	<b>Date</b>	<b>Author</b>	<b>Reason for change</b>
Version 1	1 April 2005	Kate Riggir	Document into NDHA template
Version 2	5 April 2005	Kate Riggir	Addition of glossary terms
Version 3	15 July 2005	Kirsty Smith	Amendment to Appendix C

## Table of Contents

<b>APPENDIX A - GLOSSARY .....</b>	<b>4</b>
<b>APPENDIX B - SIZING .....</b>	<b>9</b>
<b>APPENDIX C - EXISTING SOFTWARE AND HARDWARE INFRASTRUCTURE .....</b>	<b>10</b>
1.1.    SERVERS .....	10
1.2.    APPLICATIONS .....	10
1.3.    LAN AND WAN .....	13
1.4.    WORKSTATIONS .....	14
1.5.    DESKTOP .....	14

## Appendix A - Glossary

<b>Access</b>	The OAIS entity that contains the services and functions which make the archival information holdings and related services visible to Consumers.	CCSDS Blue Book <sup>1</sup>
<b>Access Method</b>	A method for retrieving an Archival Information Package based on its name or identifier, which is available to authorized users.	CCSDS Blue Book
<b>Acquisition</b>	The process of selecting, ordering, and receiving materials for library or archival collections by purchase, exchange, or gift, which may include budgeting and negotiating with outside agencies, such as publishers, dealers, and vendors, to obtain resources to meet the needs of the institution's clientele, in the most economical and expeditious manner.	ODLIS: Online Dictionary for Library and Information Science
<b>Adhoc Order</b>	A request that is generated by a Consumer for information the OAIS has indicated is currently available.	CCSDS Blue Book
<b>Administration</b>	The OAIS entity that contains the services and functions needed to control the operation of the other OAIS functional entities on a day-to-day basis.	CCSDS Blue Book
<b>AIP</b>	Archival Information Package. The Library considers an AIP to be the digital object and all of its associated metadata.	NLNZ
<b>Archival Storage</b>	The OAIS entity that contains the services and functions used for the storage and retrieval of Archival Information Packages.	CCSDS Blue Book
<b>Archive</b>	An organization that intends to preserve information for access and use by a Designated Community.	CCSDS Blue Book
<b>Client</b>	An application that exchanges information with another application (see also Consumer).	CCSDS Blue Book
<b>Consumer</b>	The role played by those persons, or client systems, who interact with OAIS services to find preserved information of interest and to access that information in detail. This can include other OAISs, as well as internal OAIS persons or systems.	CCSDS Blue Book
<b>Content Information</b>	The set of information that is the original target of preservation. It is an Information Object comprised of its Content Data Object and its Representation Information. An example of Content Information could be a single table of numbers representing, and understandable as, temperatures, but excluding the documentation that would explain its history and origin, how it relates to other observations, etc.	CCSDS Blue Book
<b>Data</b>	A reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing. Examples of data include a sequence of bits, a table of numbers, the characters on a page, the recording of sounds made by a person speaking, or a moon rock specimen.	CCSDS Blue Book

<sup>1</sup> Consultative Committee for Space Data Systems (2002), "Reference Model for an Open Archival Information System (OAIS)", CCSDA 605.0-R-1 – Blue Book.

<b>Data Dictionary</b>	A formal repository of terms used to describe data.	CCSDS Blue Book
<b>Data Management</b>	The OAIS entity that contains the services and functions for populating, maintaining, and accessing a wide variety of information. Some examples of this information are catalogs and inventories on what may be retrieved from Archival Storage, processing algorithms that may be run on retrieved data, Consumer access statistics, Consumer billing, Event Based Orders, security controls, and OAIS schedules, policies, and procedures.	CCSDS Blue Book
<b>Data Management Data</b>	The data created and stored in Data Management persistent storage that refer to operation of an archive. Some examples of this data are accounting data for Consumer billing and authorization, policy data, Event Based Order (subscription) data for repeating requests, preservation process history data, and statistical data for generating reports to archive management.	CCSDS Blue Book
<b>Data Object</b>	Either a Physical Object or a Digital Object.	CCSDS Blue Book
<b>Descriptive Information</b>	The set of information, consisting primarily of Package Descriptions, which is provided to Data Management to support the finding, ordering, and retrieving of OAIS information holdings by Consumers	CCSDS Blue Book
<b>Designated Community</b>	An identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities.	CCSDS Blue Book
<b>DIDL</b>	Digital Item Declaration Language DIDL is contained in MPEG-21.	NLNZ
<b>Digital Object</b>	An object composed of a set of bit sequences. A single intellectual entity. At any one time the Library may hold more than one "instance" of an object. In turn, an individual instance may consist of one or more files. For each object the Library will record and maintain a set of metadata. Metadata will be held for the object as a whole, each instance of the object and each part of an instance.	NLNZ
<b>Digital Objects</b>	Digital objects encompasses both "born digital" material, i.e. materials which have no analogue equivalent, and digital surrogates created as a result of converting analogue material to digital form.	NLNZ
<b>DIP</b>	Dissemination Information Package. The Library has interpreted the term to be: One or more digital objects and associated metadata or one or more items of metadata.	NLNZ
<b>E-publications</b>	Electronic publications.	NLNZ
<b>Event Based Order</b>	A request that is generated by a Consumer for information that is to be delivered periodically on the basis of some event or events.	CCSDS Blue Book

<b>Finding Aid</b>	A type of Access Aid that allows a user to search for and identify Archival Information Packages of interest.	CCSDS Blue Book
<b>Fixity</b>	Processes that confirm the integrity of a file, typically used to ensure no corruption has occurred to a file, e.g. as a result of processes such as data transmission or virus checks.	PREMIS
<b>Fixity Information</b>	The information that documents the authentication mechanisms and provides authentication keys to ensure that the Content Information object has not been altered in an undocumented manner. An example is a Cyclical Redundancy Check (CRC) code for a file.	CCSDS Blue Book
<b>Information</b>	Any type of knowledge that can be exchanged. In an exchange, it is represented by data. An example is a string of bits (the data) accompanied by a description of how to interpret a string of bits as numbers representing temperature observations measured in degrees Celsius (the representation information).	CCSDS Blue Book
<b>Information Package</b>	The Content Information and associated Preservation Description Information that is needed to aid in the preservation of the Content Information. The Information Package has associated Packaging Information used to delimit and identify the Content Information and Preservation Description Information.	CCSDS Blue Book
<b>Ingest</b>	The services and functions to: Accept digital material from Producers, Perform quality assurance checks on the digital material, Reject material that does not meet quality standards Reject material that does not fall within the Library's collection policy, Prepare accepted digital material for storage and management, Create and derive metadata required for the coordination and management of information within the NDHA.	NLNZ
<b>Knowledge Base</b>	A set of information, incorporated by a person or system, that allows that person or system to understand received information.	CCSDS Blue Book
<b>Long Term</b>	A period of time long enough for there to be concern about the impacts of changing technologies, including support for new media and data formats, and of a changing user community, on the information being held in a repository. This period extends into the indefinite future.	CCSDS Blue Book
<b>Long Term Preservation</b>	The act of maintaining information, in a correct and Independently Understandable form, over the Long Term.	CCSDS Blue Book
<b>Management</b>	The role played by those who set overall OAIS policy as one component in a broader policy domain.	CCSDS Blue Book

<b>Metadata</b>	Data about other data	CCSDS Blue Book
<b>METS</b>	Metadata Encoding and Transmission Standard. The METS schema is a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, expressed using the XML schema language of the World Wide Web Consortium. The standard is maintained in the Network Development and MARC Standards Office of the Library of Congress, and is being developed as an initiative of the Digital Library Federation.	NLNZ
<b>MPEG-21</b>	<p>MPEG-21 is a Multimedia Framework that aims at defining a normative open framework for multimedia delivery and consumption for use by all the players in the delivery and consumption chain. This open framework will provide content creators, producers, distributors and service providers with equal opportunities in the MPEG-21 enabled open market. This will also be to the benefit of the content consumer providing them access to a large variety of content in an interoperable manner</p> <p>MPEG-21 is based on two essential concepts: the definition of a fundamental unit of distribution and transaction (the Digital Item) and the concept of Users interacting with Digital Items. The Digital Items can be considered the “what” of the Multimedia Framework (e.g., a video collection, a music album) and the Users can be considered the “who” of the Multimedia Framework.</p> <p>The goal of MPEG-21 can thus be rephrased to: defining the technology needed to support Users to exchange, access, consume, trade and otherwise manipulate Digital Items in an efficient, transparent and interoperable way.</p> <p><a href="http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm">http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm</a></p>	NLNZ
<b>NDHA</b>	National Digital Heritage Archive.	NLNZ
<b>OAIS</b>	Open Archival Information System. The Reference Model for an OAIS is a conceptual framework for an archival system dedicated to preserving and maintaining access to digital information over the long term. RLG, OCLC, many members of our respective organizations, and several groundbreaking international projects have played key roles in shaping the reference model and in adapting the reference model for use in libraries, archives and research repositories. And yet libraries and archives represent only a section of the community with interest in the OAIS reference model.	CCSDS Blue Book
<b>PDI</b>	Preservation Description Information. The information that is necessary for preservation of digital objects.	NLNZ
<b>PREMIS</b>	PREservation Metadata: Implementation Strategies. A working group aimed at the development of recommendations and best practices for implementing preservation metadata. Sponsored by the OCLC Office of Research, OCLC Member	NLNZ

	Services, and RLG.	
<b>Producer</b>	The role played by those persons, or client systems, who provide the information to be preserved. This can include other OAISs or internal OAIS persons or systems.	CCSDS Blue Book
<b>Reference Model</b>	A framework for understanding significant relationships among the entities of some environment, and for the development of consistent standards or specifications supporting that environment. A reference model is based on a small number of unifying concepts and may be used as a basis for education and explaining standards to a non-specialist.	CCSDS Blue Book
<b>Refreshment</b>	A Digital Migration where the effect is to replace a media instance with a copy that is sufficiently exact that all Archival Storage hardware and software continues to run as before.	CCSDS Blue Book
<b>Repackaging</b>	A Digital Migration in which there is an alteration in the Packaging Information of the AIP.	CCSDS Blue Book
<b>Representation Information</b>	The information that maps a Data Object into more meaningful concepts. An example is the ASCII definition that describes how a sequence of bits (i.e., a Data Object) is mapped into a symbol.	CCSDS Blue Book
<b>Result Set</b>	The set of descriptive records for those AIPs in an OAIS that match the criteria stated in a Consumer query, or other results from a search on Data Management.	CCSDS Blue Book
<b>SIP</b>	Submission Information Package	CCSDS Blue Book
<b>Staging Area</b>	A work space within the NDHA as opposed to an archival storage place.	NLNZ
<b>Submission Agreement</b>	The agreement reached between an OAIS and the Producer that specifies a data model for the Data Submission Session. This data model identifies format/contents and the logical constructs used by the Producer and how they are represented on each media delivery or in a telecommunication session.	CCSDS Blue Book
<b>Transformation</b>	A Digital Migration in which there is an alteration to the Content Information or PDI of an Archival Information Package. For example, changing ASCII codes to UNICODE in a text document being preserved is a Transformation.	CCSDS Blue Book
<b>Version</b>	An attribute of an AIP whose information content has undergone a transformation on a source AIP and is a candidate to replace the source AIP.	CCSDS Blue Book

## Appendix B - Sizing

Item	Source	Digitisation	Web Harvesting	Sound (.BWF)	Unpublished	e-Legal
Volumes  (Digital objects only, excludes associated metadata)	Totals  2005 – 9,534 Gb 2006 – 17,138Gb 2007 – 24,752 Gb 2008 – 32,391 Gb 2009 – 40,076 Gb	Based on current staffing levels (2.8 FTE) 10500 items pa at a total size of 1487 Gig pa  Based on sought after staffing levels (3.8 FTE) 16100 items pa at a total size of 2871 Gig pa  NB Disproportionate increase in size compared with number of items is because of additional items having a higher % of coloured items.	Based on 1.5FTE to do the work:  2005/06 1800 sites @ 600 Mb each = 1080 Gig  2006/07 3600 sites @ 600 Mb each = 2160 Gig  2007/08 5400 sites @ 600 Mb each = 3240 Gig	Est 600 per annum non jukebox @ 1.5Gig per item = 900 Gig pa  Est 200 juke box per annum @1.5 Gig per item = 300 Gig pa	Based on 100% increase in new objects each year	Based on 20,000 objects within 5 years at an average size of 20Mb
Cumulative estimates		2005 – 7153Gb 2006 – 12391Gb 2007 – 17633 Gb 2008 – 22888G 2009 – 28166G  Based on 3.8 FTEs and the same volumes each year	2005 – 1080 Gb 2006 – 2160 Gb 2007 – 3240 Gb 2008 – 4320 Gb 2009 – 5400 Gb  Volumes assume an additional 1800 sites harvested each year plus sites from previous year.	2005 – 1200 Gb 2006 – 2400 Gb 2007 – 3600 Gb 2008 – 4800 Gb 2009 – 6000 Gb  Based on same volumes each year	2005 = 20.90 Gb 2006 = 26.85 Gb 2007 = 38.74 Gb 2008 = 62.52 Gb 2009 = 110.08 Gb	2005 – 80 Gb 2006 – 160 Gb 2007 – 240 Gb 2008 – 320 Gb 2009 – 400 Gb

## Appendix C - Existing Software and Hardware Infrastructure

The following sections detail the Library's current technical infrastructure. The current architecture should be regarded as the defacto standard for the library.

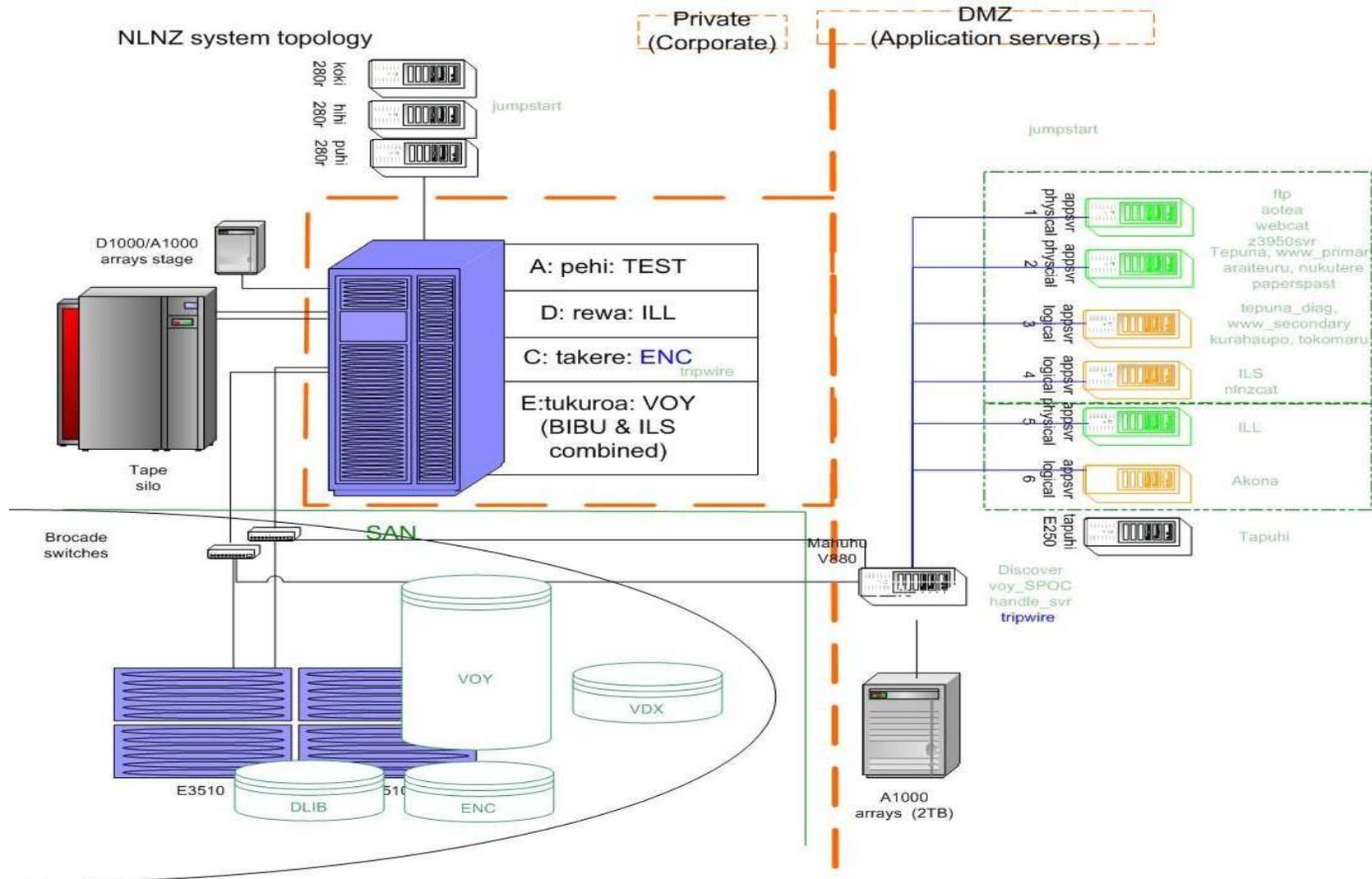
### 1.1. Servers

Hardware	Type
Servers	Enterprise class Sun servers Including Sun Fire F12K.
Operating System	Solaris 8 Solaris 9
Database	Oracle 9i Oracle 10

### 1.2. Applications

Application	Description
National Bibliographic Database (Te Puna)	Software: Voyager library management system, v2001.2 <a href="http://www.endinfosys.com/prods/voyager.htm">http://www.endinfosys.com/prods/voyager.htm</a> Hardware: Sun Fire 12K 4 x Sun E450
Integrated Library System	Software: Voyager library management system, v2001.2 <a href="http://www.endinfosys.com/prods/voyager.htm">http://www.endinfosys.com/prods/voyager.htm</a> Hardware: Sun Fire 12K Sun E450
Various Online Collections	Software: ENCompass for Digital Collections & ENCompass for Resource Access v3.5.1 LinkFinderPlus v4 <a href="http://encompass.endinfosys.com/">http://encompass.endinfosys.com/</a> Hardware: Sun V880

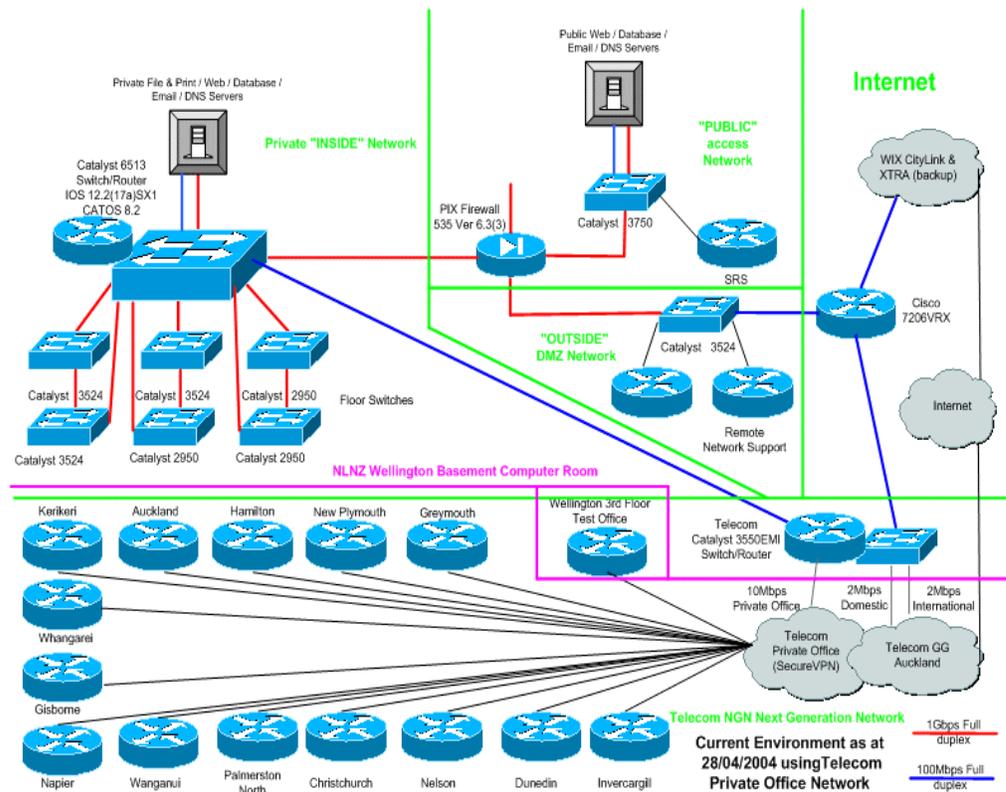
Application	Description
TAPUHI	<p>Software:</p> <p>The TAPUHI system is based on the AWAIRS product from Sanderson Computers Ltd. AWAIRS is a text based proprietary library system that uses Unidata as its database system. TAPUHI incorporates a product called SPYDUS that provides a Web like front end, using TCP/IP protocols, for both staff and public users to access collections. SPYDUS is used for read only functionality. All updates are performed using character-based screens.</p> <p>TAPUHI has been designed specifically for the Library and has been in operation since 1992. Although versions of the product have been implemented at two other major Australasian research libraries (the Hocken in Dunedin and the Mitchell in Sydney) and other versions of AWAIRS are used in a variety of other heritage institutions, it is best regarded as a commissioned development for ATL rather than as a packaged product.</p> <p>Hardware:</p> <p style="padding-left: 40px;">Sun E250</p>
National Inter Library Loans	<p>Software:</p> <p style="padding-left: 40px;">VDX</p> <p style="padding-left: 40px;"><a href="http://www.fdusa.com/products/vdx.html">http://www.fdusa.com/products/vdx.html</a></p> <p>Hardware:</p> <p style="padding-left: 40px;">Sun Fire 12K</p> <p style="padding-left: 40px;">Sun E450</p>



### 1.3. LAN and Wan

The Private 'INSIDE' Network section of the diagram on the following page describes the Library's corporate local and wide area network infrastructure.

Network	Bandwidth
WAN	100mbs
LAN	100mbs, 1gbs
Cabling	CAT5



### 1.4. Workstations

Capacity	49% (275)	32% (176)	19% (105)
Model (Compaq/HP)	EVO D510	EVO D530	EVO D530
Processor	Pentium P4	Pentium P4	Pentium P4
Speed	1.8Ghz	3Ghz	3.2Ghz
RAM	512mb	512mb	1024mb
Disk Capacity	40gb	80gb	80gb

### 1.5. Desktop

Application	Software
Operating System	Windows 2000 professional SP4
Office Suite	MS Office 2000 SP3
Mail	GroupWise 6.5
Browser	Internet Explorer 6.0