Handle System Overview and Update

Larry Lannom

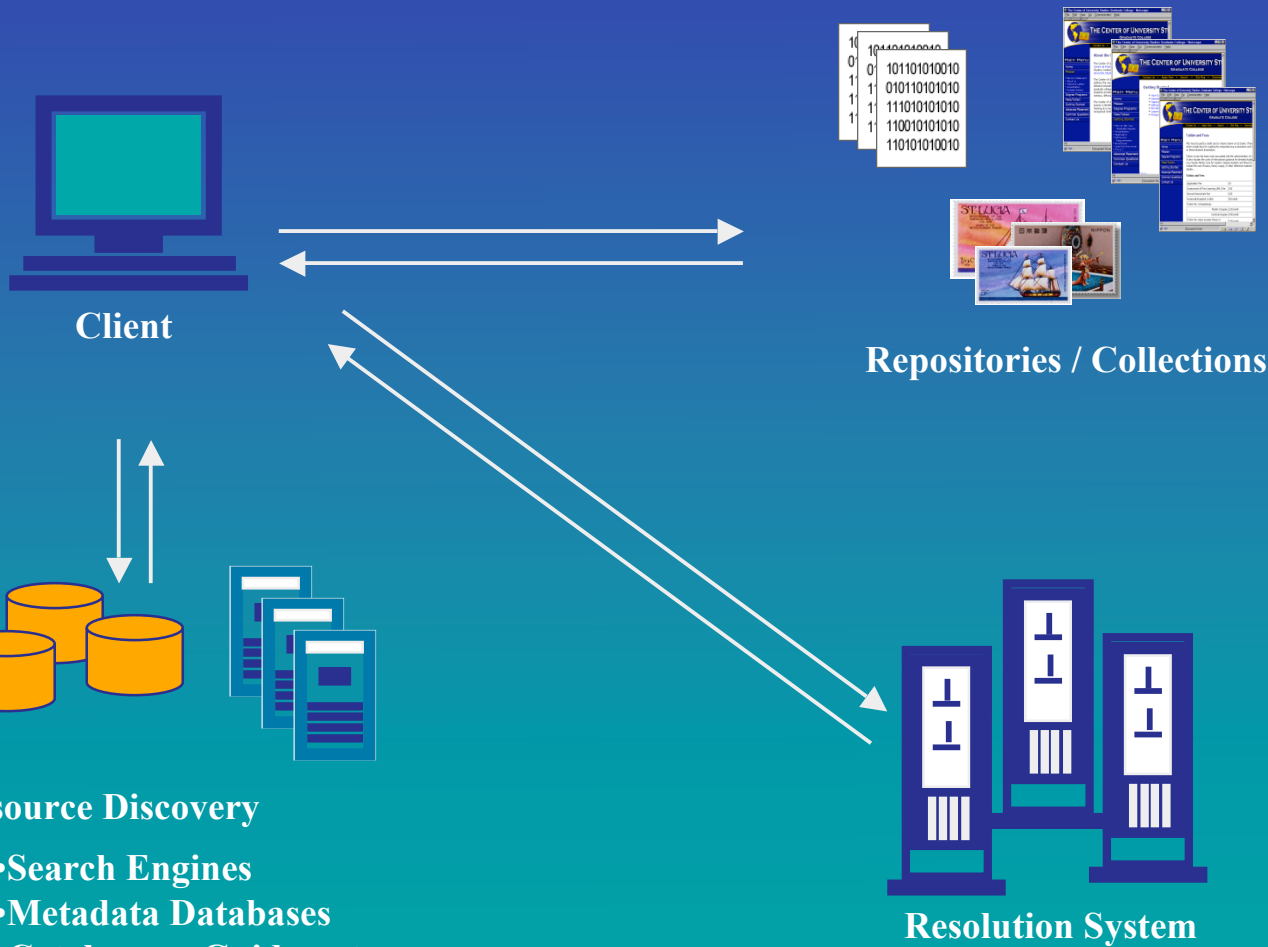
27 March 2008

Corporation for National Research Initiatives
Reston, VA
http://www.cnri.reston.va.us/
http://www.handle.net/

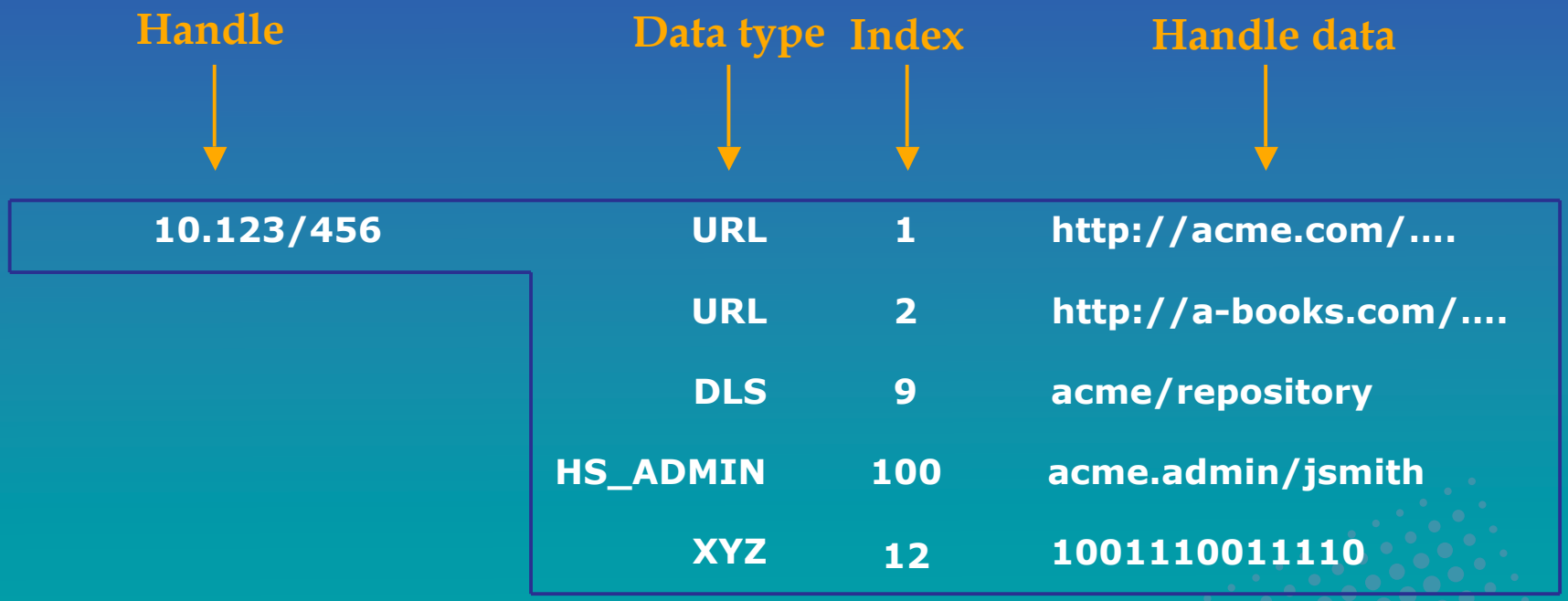
Handle System

- Provides basic identifier resolution system for Internet
  - Go from object name to current state data
  - Name can persist over changes in location and other attributes
- Logically centralized, but physically distributed and highly scalable
- Enables association of one or more typed values, e.g., IP address, public key, URL, with each id
- Optimized for resolution speed and reliability
- Secure resolution with its own PKI as an option
- Open, well-defined protocol and data model
- Provides infrastructure for application domains, e.g., digital libraries & publishing, data mgmt, network mgmt, id mgmt ...

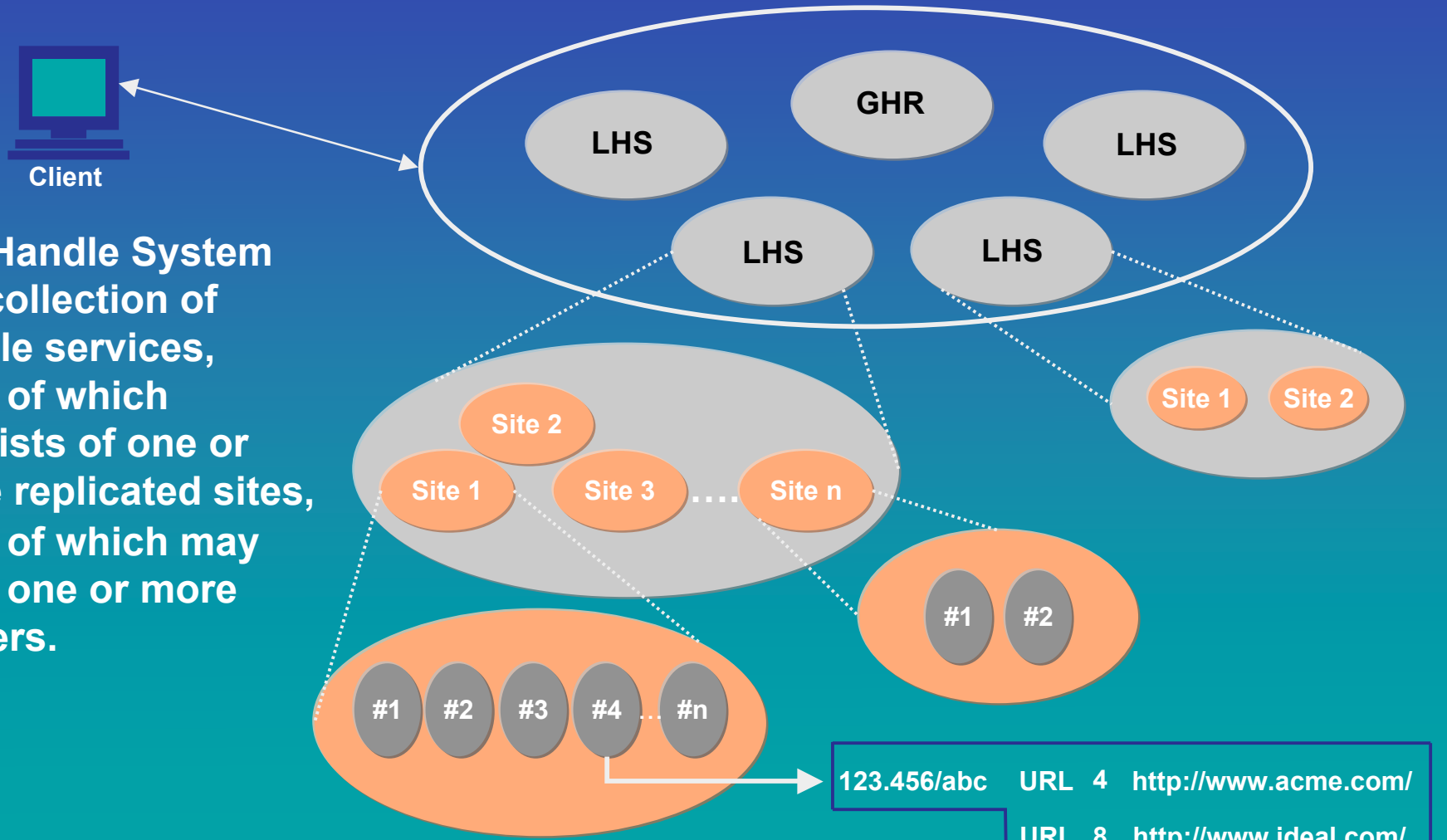
Role of Resolution System in Information Management on Networks



Handles Resolve to Typed Data



Handle Resolution



The Handle System is a collection of handle services, each of which consists of one or more replicated sites, each of which may have one or more servers.

Handle Clients

Request to Client:
Resolve hdl:10.1000/1

Client

1. Sends request to Global to resolve 0.NA/10.1000 (naming authority handle for 10.1000)

Global Handle Registry

Handle Clients

Request to Client:
Resolve hdl:10.1000/1

Client

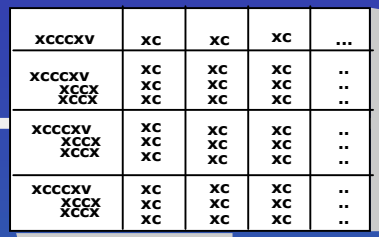
2. Global Responds with
Service Information for 10.1000

Global Handle
Registry

xccccxv	xc	xc	xc	...
xccccxv xccx xccx	xc xc xc	xc xc xc	xc xc xc	.. .. ..
xccccxv xccx xccx	xc xc xc	xc xc xc	xc xc xc	.. .. ..
xccccxv xccx xccx	xc xc xc	xc xc xc	xc xc xc	.. .. ..

Service Information
Acme Local Handle Service

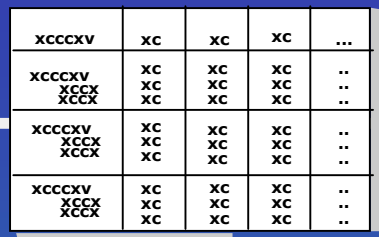
Handle Clients



	IP Address	Port #	Public Key	…
Primary Site				
Server 1	**123.45.67.8**	**2641**	**K03RLQ…**	**…**
Server 2	**123.52.67.9**	**2641**	**5&M#FG…**	**…**
Secondary Site A				
Server 1	**321.54.678.12**	**2641**	**F^*JLS…**	**…**
Server 2	**321.54.678.14**	**2641**	**3E$T%…**	**…**
Server 3	**762.34.1.1**	**2641**	**A2S4D…**	**…**
Secondary Site B				
Server 1	**123.45.67.4**	**2641**	**N0L8H7…**	**…**

Service Information - Acme Local Handle Service

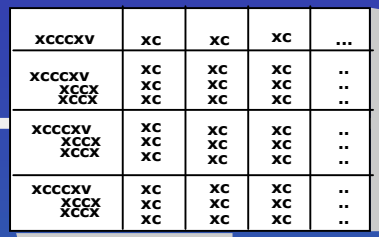
Handle Clients



	IP Address	Port #	Public Key	…
Primary Site				
Server 1	123.45.67.8	2641	K03RLQ…	…
Server 2	123.52.67.9	2641	5&M#FG…	…
Secondary Site A				
Server 1	321.54.678.12	2641	F^*JLS…	…
Server 2	321.54.678.14	2641	3E$T%…	…
Server 3	762.34.1.1	2641	A2S4D…	…
Secondary Site B				
Server 1	123.45.67.4	2641	N0L8H7…	…

Service Information - Acme Local Handle Service

Handle Clients

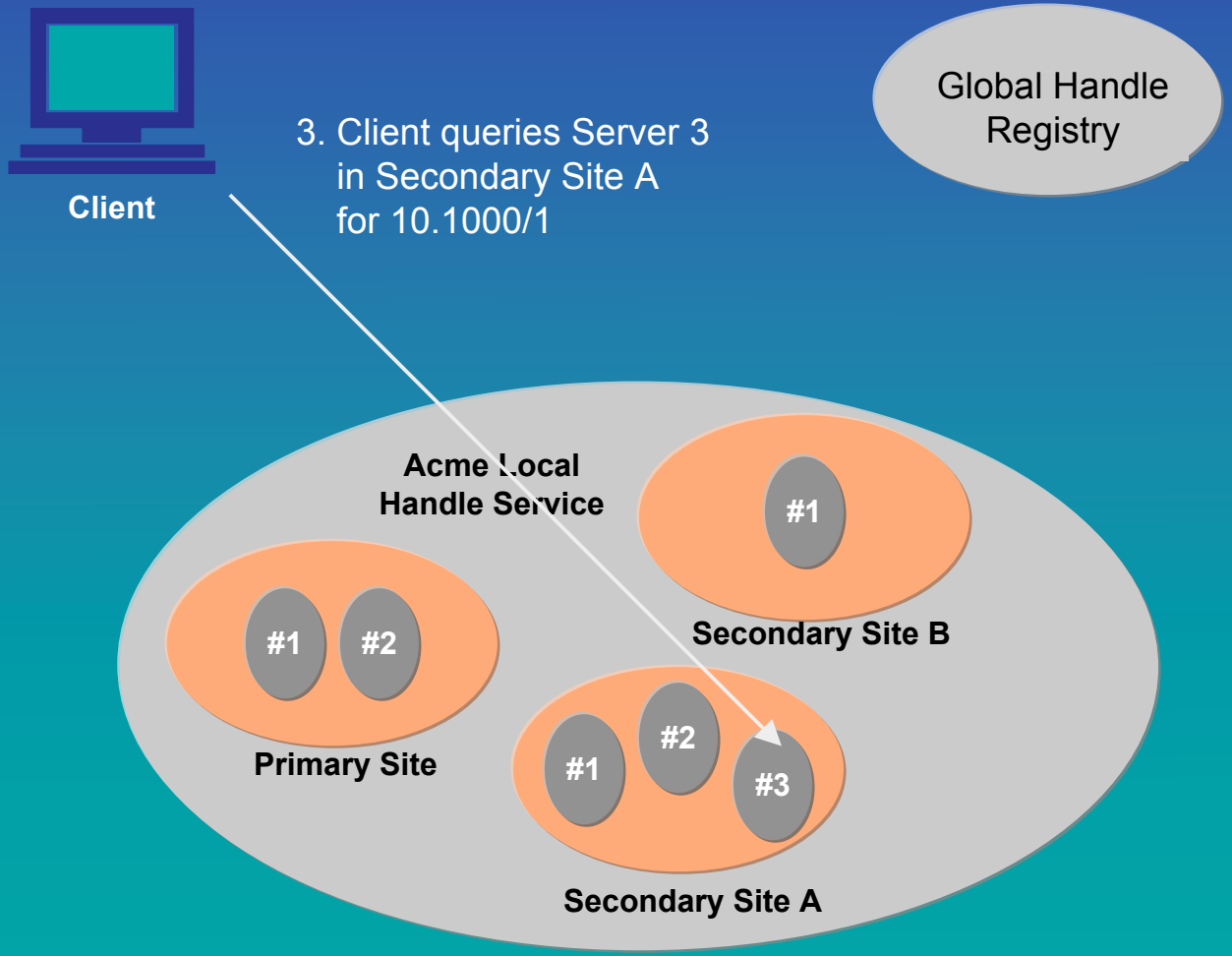


	IP Address	Port #	Public Key	…
Primary Site				
Server 1	123.45.67.8	2641	K03RLQ…	…
Server 2	123.52.67.9	2641	5&M#FG…	…
Secondary Site A				
Server 1	321.54.678.12	2641	F^*JLS…	…
Server 2	321.54.678.14	2641	3E$T%…	…
Server 3	762.34.1.1	2641	A2S4D…	…
Secondary Site B				
Server 1	123.45.67.4	2641	N0L8H7…	…

Service Information - Acme Local Handle Service

Handle Clients

Request to Client:
Resolve hdl:10.1000/1



Handle Clients

Request to Client:
Resolve hdl:10.1000/1

Client

Global Handle Registry

4. Server responds with handle data

Acme Local Handle Service

#1

Secondary Site B

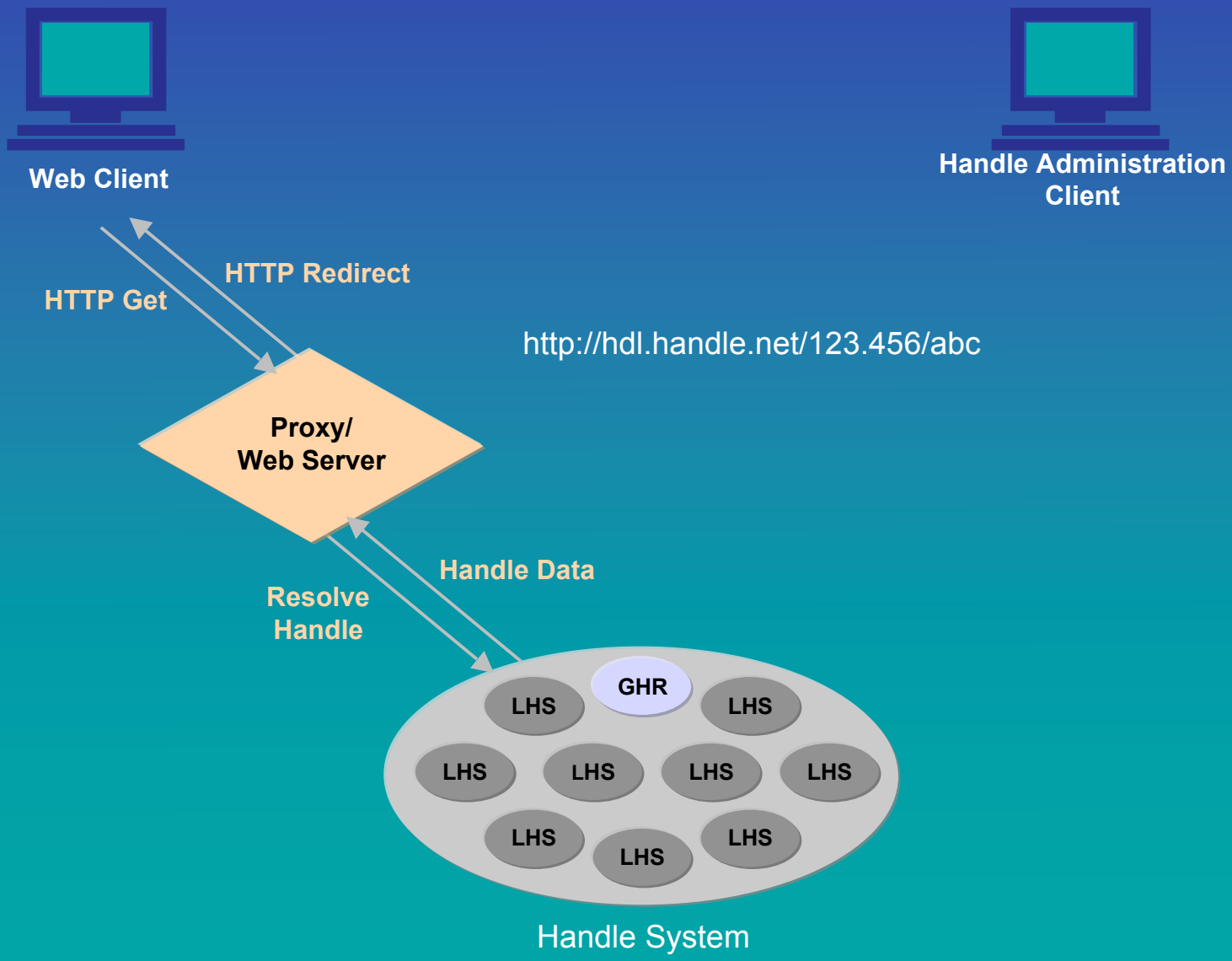
#1 #2

Primary Site

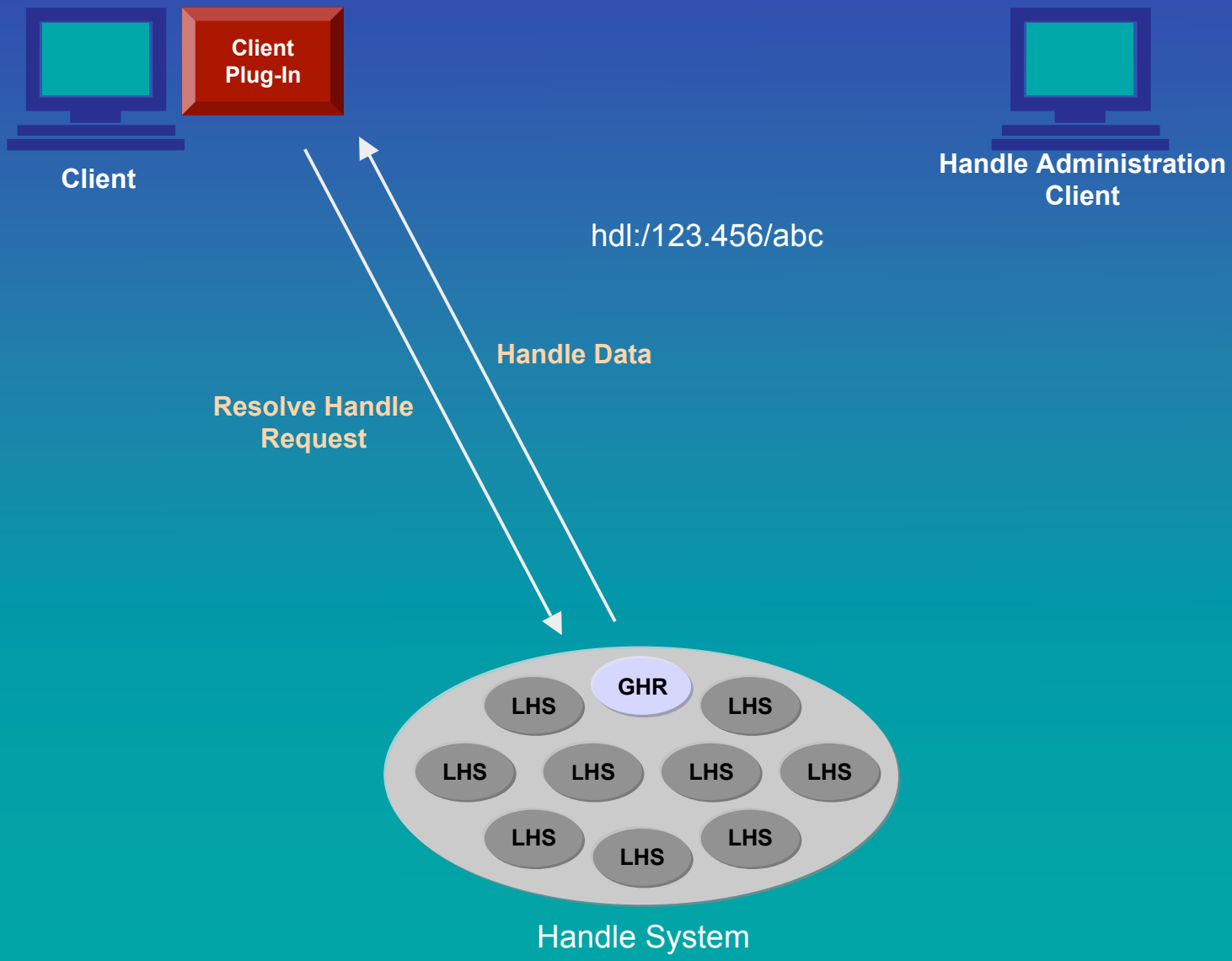
#1 #2 #3

Secondary Site A

Handle Clients



Handle Clients



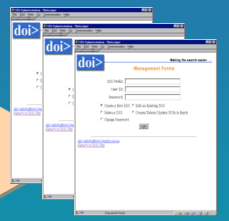
Handle Clients

Web

Handle Administration Client

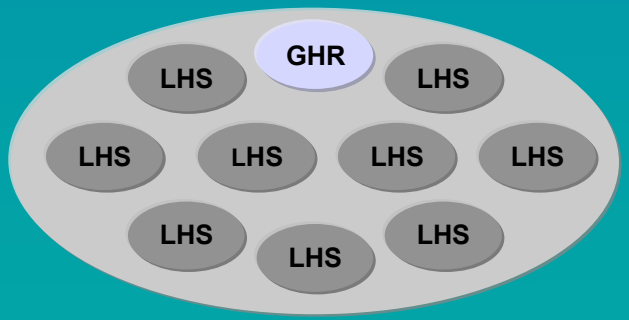
HTTP

Web Server



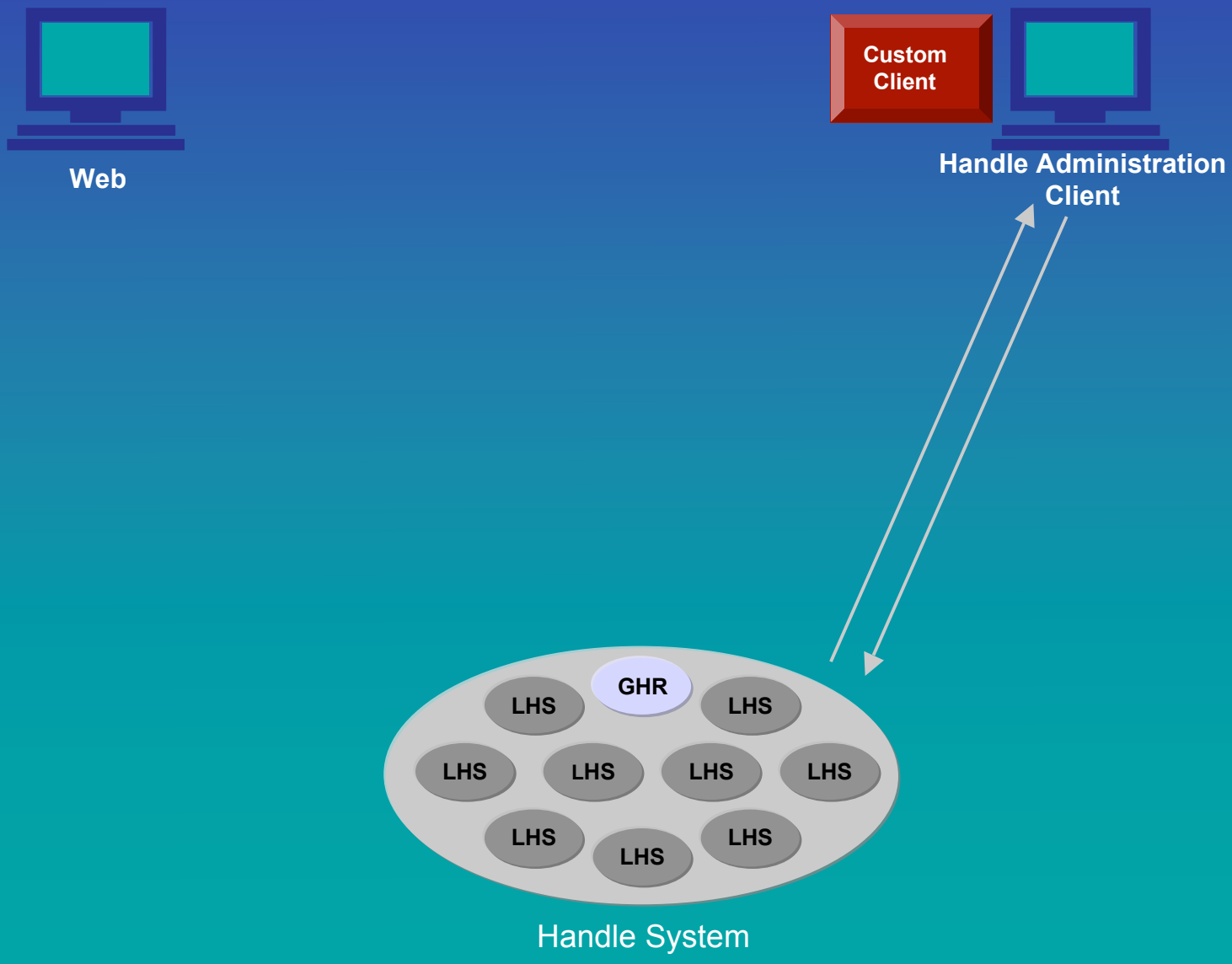
Admin Forms

Handle Admin API



Handle System

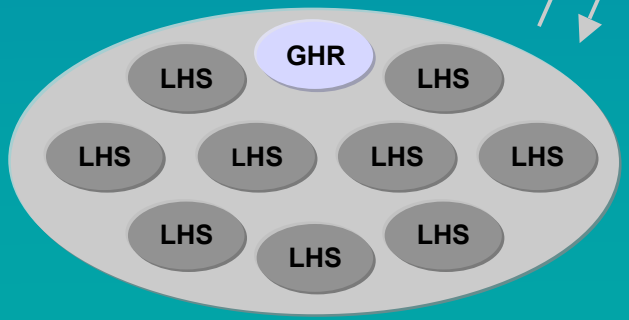
Handle Clients



Handle Clients

Web

Handle Administration embedded in another process

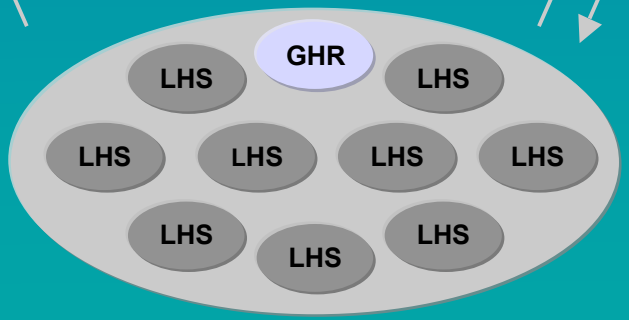


Handle System

Handle Clients

Handle Resolution embedded in another process

Handle Administration embedded in another process



Handle System

Handle System Usage

- Library of Congress
- DTIC (Defense Technical Information Center)
- IDF (International DOI Foundation)
  - CrossRef (scholarly journal consortium, representing >2K publishers & societies)
  - CAL (Copyright Agency Ltd - Australia)
  - MEDRA (Multilingual European DOI Registration Agency)
  - Nielsen BookData (bibliographic data - ISBN)
  - R.R. Bowker (bibliographic data - ISBN)
  - Office of Publications of the European Community (OPOCE)
  - German National Library of Science and Technology (TIB)
  - Wanfang Data
- OECD
- NASA
- National Agricultural Library/USDA
- DSpace (MIT + HP)
- ADL (DoD Advanced Distributed Learning initiative)
- Los Alamos National Laboratory Research Library
- Australian Dept. of Ed., Sci, and Training (DEST) - PILIN project
- Clarin (Common Language Resources and Technology Infrastructure)

Handle System Usage
As of 1 March 08

- Assigned Prefixes
  - DOI - 2,560
  - Other - 1,135
- Handles
  - DOI - 33M +
  - Other - Additional millions (total per prefix known only to prefix manager; LANL adding 600M but privately)
- Handle Services
  - Global
    - Core: three service sites (added locations being worked)
    - ~100M resolutions per month
  - Locals
    - ~ 1000 registered LHS's

Handle System Management and Standards

- Specification
  - RFC 3650: Overview
  - RFC 3651: Namespace and Service Definition
  - RFC 3652: Protocol
- DoDI 1322
- ISO standards track for DOI
- HSAC - Handle System Advisory Committee
  - Approx 15 members representing big users
  - Goal: evolve to oversee the system

Handle System Public Licensing

- License
  - HS Version 6.2 released June 2006 under public license
  - Commercial use welcomed
    - no longer restricted to research and/or education
  - No licensing fees for software or underlying technology
- Service Agreement
  - Service Agreement is required if you use the software/underlying technology to resolve identifiers
  - One time $50 registration fee per prefix
  - Annual $50 maintenance fee per prefix\
  - Five and ten year discounts available
  - Fees needed to help support global root and administration

New and Upcoming Technical Developments "Chooseby"

- Structured alternatives, e.g., multiple locations, in a single handle value
- Include selection criteria in that same value
- Handle client application, e.g., proxy server, performs evaluation
- Type = 10320/loc; value =

  <locations chooseby locatt, country,weight>

  <Location id=0 href="http://abc…. Country="gb" weight=0>

  <Location id=1 href="http://def… weight=1>

  <Location id=2 href="http://xyz… weight=1>

  <locations/>
- Prototyped as part of proxy upgrade
- Will be deployed in dx.doi.org, and will be part of generic proxy download
- Approach extensible for future selection methods, e.g., chooseby language

New and Upcoming Technical Developments
Type Registry

- Each *handle value* is typed to define syntax and semantics of the data
- Type strings are themselves handles, either explicitly or implicity
- Implicit handles are 0.Type handles, e.g., 0.Type/HS_ADMIN
- Many handle application developers create their own types
- Good news/Bad news
  - No constraints on type creation
  - No standard way to share types or properly interpret the types of others
- Solution: optional structured description plus public type registry
- Structured description
  - XML encoded set of standard values, e.g., POC, natural language description, encoding spec, external reference documents, rendering proxies, etc.
  - Values in-line or by reference
  - Structured description would be one of the type/value pairs for the type handle
- Public Type Registry
  - Register existing types to encourage re-use and standardization
  - Search the simple descriptions and, if available, structured description
  - Handle system types and CNRI project types will seed the registry
- Goal - describe this process in HaRP (Handle Recommended Practices) #1

New and Upcoming Technical Developments Delegation

- Every handle prefix is registered in the Global service under 0.NA
- Every client finds the correct LHS by searching its cache or resolving the prefix handle to get the LHS service information
- Primary advantage of relatively flat namespace - resolution speed
  - Trips on the net independent of number of prefixes and handle services
- BUT - good reasons to not want prefixes known to Global
  - You don't want Global admin to know about them
  - You want to create a lot of prefixes quickly
- Operational delegation (as opposed to organizational) prototyped but not yet deployed
  - If 0.NA prefix not found AND prefix consists of multiple segments, e.g., 0.NA/20.30.40
  - Then client will try 0.NA/20 and if "del='y'" in a certain typed value
  - Then client will follow that LHS pointer and try again
- Policy still needs to be developed ($50/prefix?)
- Anticipating, but not promising, deployment during 2008

New and Upcoming Technical Developments
Computed Handles

- Successfully resolve Handles that haven't been registered in a Handle Server
- Not really new and not a result of any specification changes
  - Highly useful, but in limited circumstances
  - Recently implemented, but not yet publicly documented, for a CNRI project. Quite similar to DSpace approach
- Directly results from the modularity of the current implementation
  - Backend handle server storage is pluggable
  - A new storage module allows handles to be computed
  - Will be made available in future releases
- Simple example (and our use case)
  - All handles starting with "123/456:" reference a specific digital object repository
  - That repository contains 10M objects, all identified as 123/456:1, 123/456:2, etc.
  - The resolution of each of those 10M handles should return the same values
  - New storage module, configured to use ":" as a delimiter, will first resolve the full handle and, if not found, will return the results for the 'stub' handle.
  - If any of the set of 10M needs unique values, that handle is separately registered
- Other computations could be performed on the string following the delimiter or on the entire handle
- Handles can remain static in reference form, while millions of resolution values can be changed at a single stroke

New and Upcoming Technical Developments Other

- More Global Mirrors
  - Currently one primary and 2 secondaries, all run by CNRI (2 physical locations and 3 network providers)
  - One new (silent) secondary run by Crossref in Boston
  - Policies in place - separate ‘no money’ deal with service guarantees on both sides
  - Monitoring procedures in test mode
  - Anticipate one to two additional by end of 2008
- Updated RFCs
  - 5 years old
  - Some of the specification never implemented in reference implementation
    - Execute permissions - needed on Handle values?
    - Delegation doesn’t look quite like we thought it would five years ago
- URI registration
  - Goal for 2008
  - Cannot speak to likelihood of acceptance
  - Info URI in place: info:hdl

PILIN Project

"Strengthening Australia's ability to use global identifier infrastructure"

1. *Develop* shared, standards-based, persistent identifier management infrastructure
2. Support *adoption* of persistent identifiers and services
3. Plan for *sustainable* shared identifier infrastructure

Supporting adoption through

- "proof of concept" deployment and demonstrator services
- documenting best practice

Timeline: October 2006 - December 2007

http://www.arrow.edu.au/PILIN.php

SOME RIGHTS RESERVED

Project outputs

1. **Model** lifecycle of persistent identifiers and services
2. Best practice and policy **guides**
3. Document community **requirements**
4. Model identifier management **services**
   - Technology-independent *service genres*
   - Technology-dependent *service expressions*
5. **Pilot** shared persistent identifier management infrastructure
   - Using Handles system
6. **Software toolkits** to aid use of the infrastructure
7. **Sustainability** options and proposals

SOME RIGHTS RESERVED