COMP90015 Distributed Systems
Semester 2, 2022
Topic: Name Services

Dr Tawfiq Islam
School of Computing and Information Systems (CIS)
The University of Melbourne, Australia

Learning Outcomes

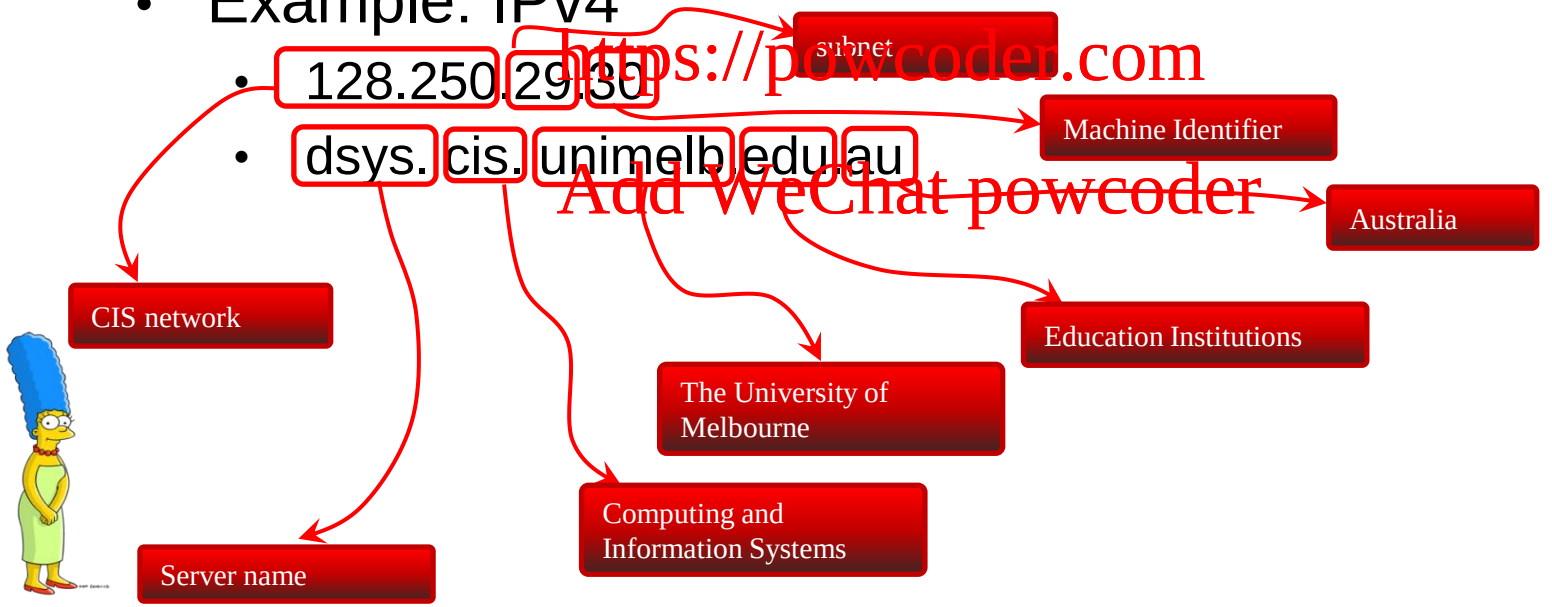
- To understand the need for naming systems in distributed systems
- To be familiar with the design requirements such as structure and management of name spaces, and operations supported by them.
- To understand the operation of the Internet naming service – DNS (Domain Name System)
- To understand structure and operation of directory services – X.500 Directory Service & LDAP (Lightweight Directory Access Protocol)
- Reading: Distributed Systems: Concepts and Design by George Coulouris (5th edition). Chapter 13. Sections: 13.1, 13.2, 13.3

Which one is easy for humans and machines? and why?

- 74.125.237.83 or google.com
- 128.250.1.25 or cis.unimelb.edu.au
- Disk 4, Sector 2, block 5 or /usr/home/tawfiq/Hello.java
- tawfiq@128.250.1.25 or tawfiqul.islam@unimelb.edu.au

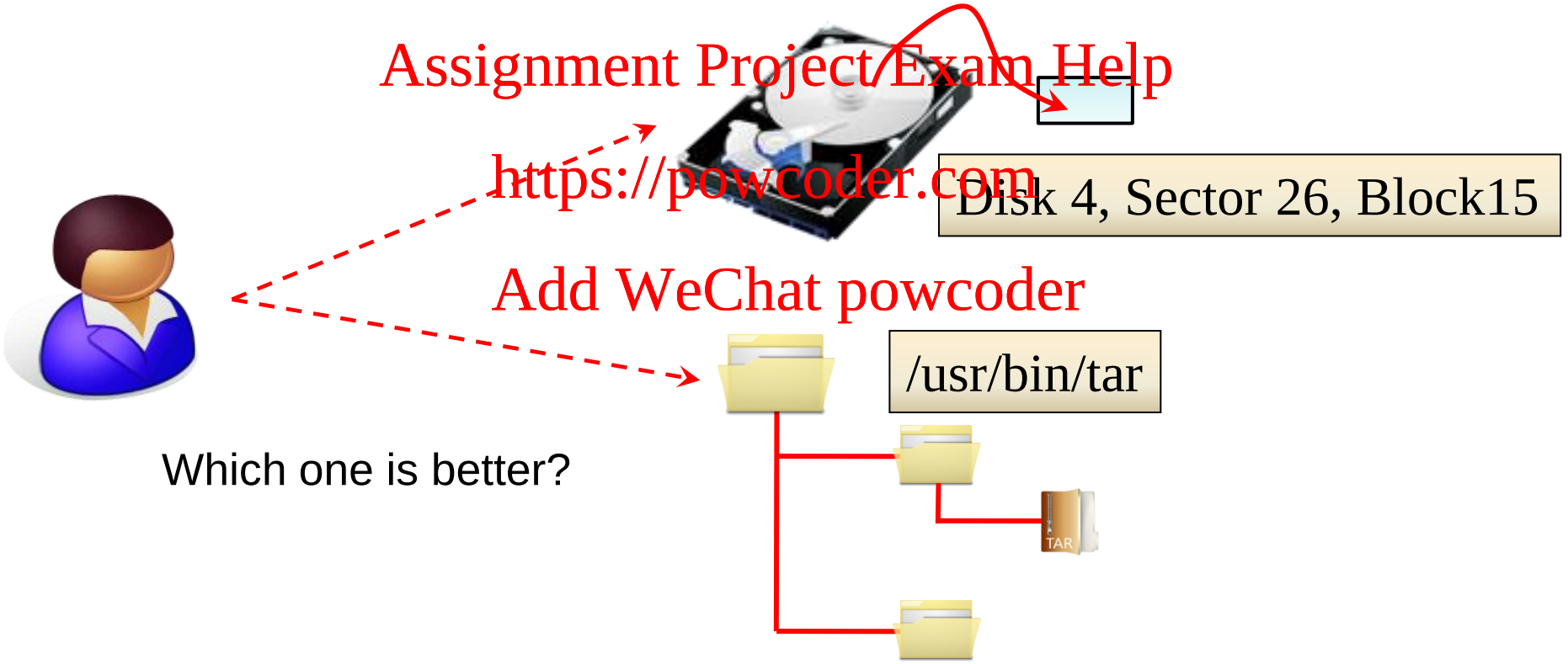
Names or Codes, or Numbers?

- Names (when meaningful) are easier to remember than codes or numbers…
- Number (or sequence codes) are more useful for structuring data and locating resources by a program..
- Example: IPv4
  - 128.250.29.30
  - dsys.cis.unimelb.edu.au



Names or Codes, or Numbers?

- As discussed in file system (hierarchical naming of files) and mounting at right location.



Which one is better?

Names in Distributed Systems

- In a distributed system, names are used to refer to a wide variety of resources such as:
  - Computers, services, remote objects, and files, as well as users.
- Naming is a fundamental issue in DS design as it facilitates communication and resource sharing.
  - A name in the form of URL is needed to access a specific web page.
  - Processes cannot share resources managed by a computer system unless they can name them consistently.
  - Users cannot communicate within one another via a DS unless they can name one another, with email address.

Naming Services

- **Definition**
  - In a Distributed System, a Naming Service is a specific service whose aim is to provide a consistent and uniform naming of resources, thus allowing other programs or services to localize them and obtain the required metadata for interacting with them.
  - A name service stores a collection of one or more naming contexts, sets of bindings between textual names and attributes for objects such as computers, services, and users.
  - The major operation that a name service supports is to resolve names.
- **Key benefits**
  - Resource localization
  - Uniform naming
  - Device independent address (e.g., you can move domain name/web site from one server to another server seamlessly).

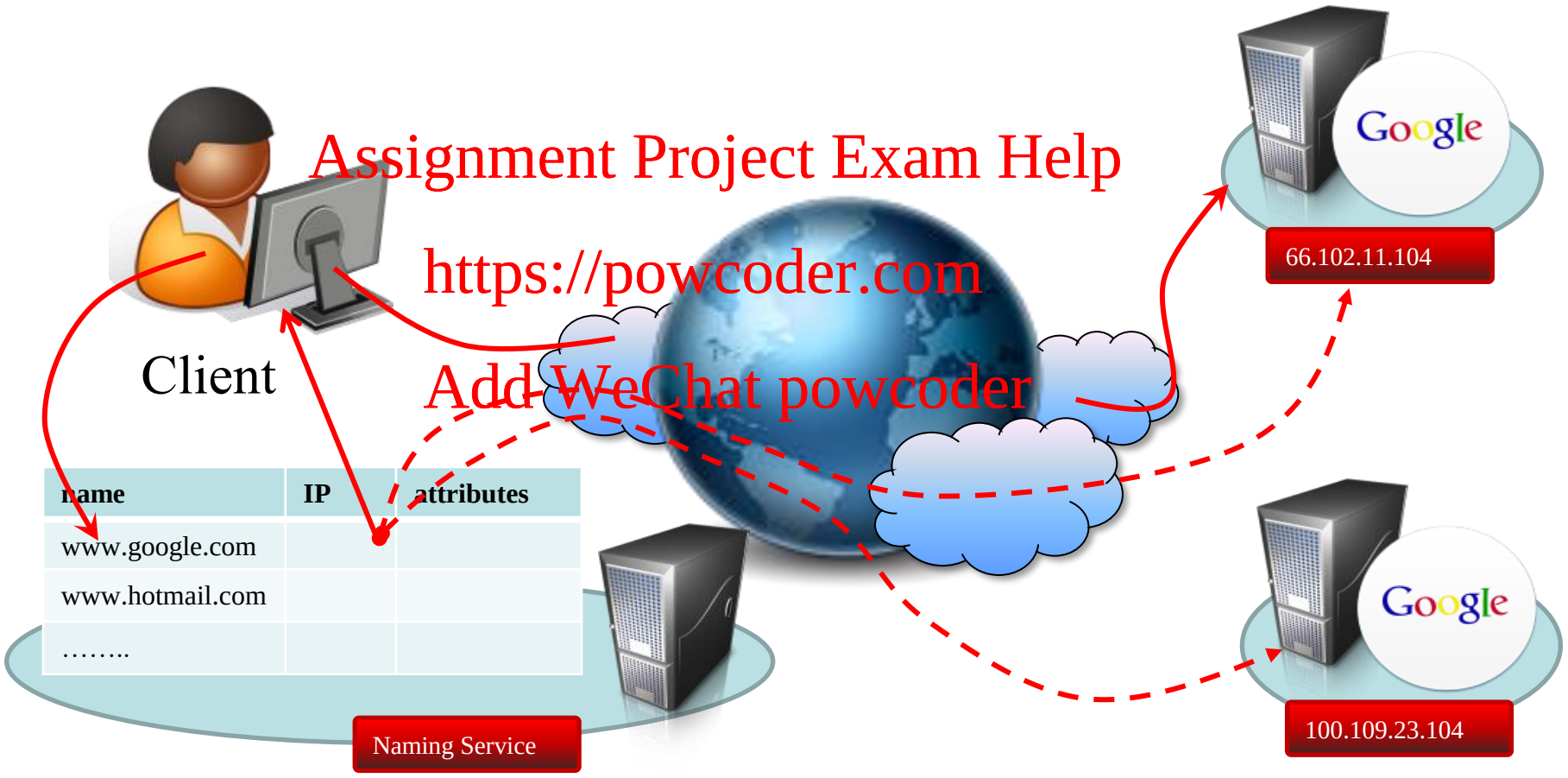
Naming Services

- How do Naming Services facilitate communication and resource sharing?
  - A URL facilitates the localization of a resource exposed on the Web.
    - e.g., abc.net.au means it is likely to be an Australian entity?
  - A consistent and uniform naming helps processes in a distributed system to interoperate and manage resources.
    - e.g., commercials use .com; non-profit organizations use .org
    - .edu, ac.uk or edu.au educational institutes
  - Users refers to each other by means of their names (i.e., email) rather than their system ids
  - Naming Services are not only useful to locate resources but also to gather additional information about them such as attributes

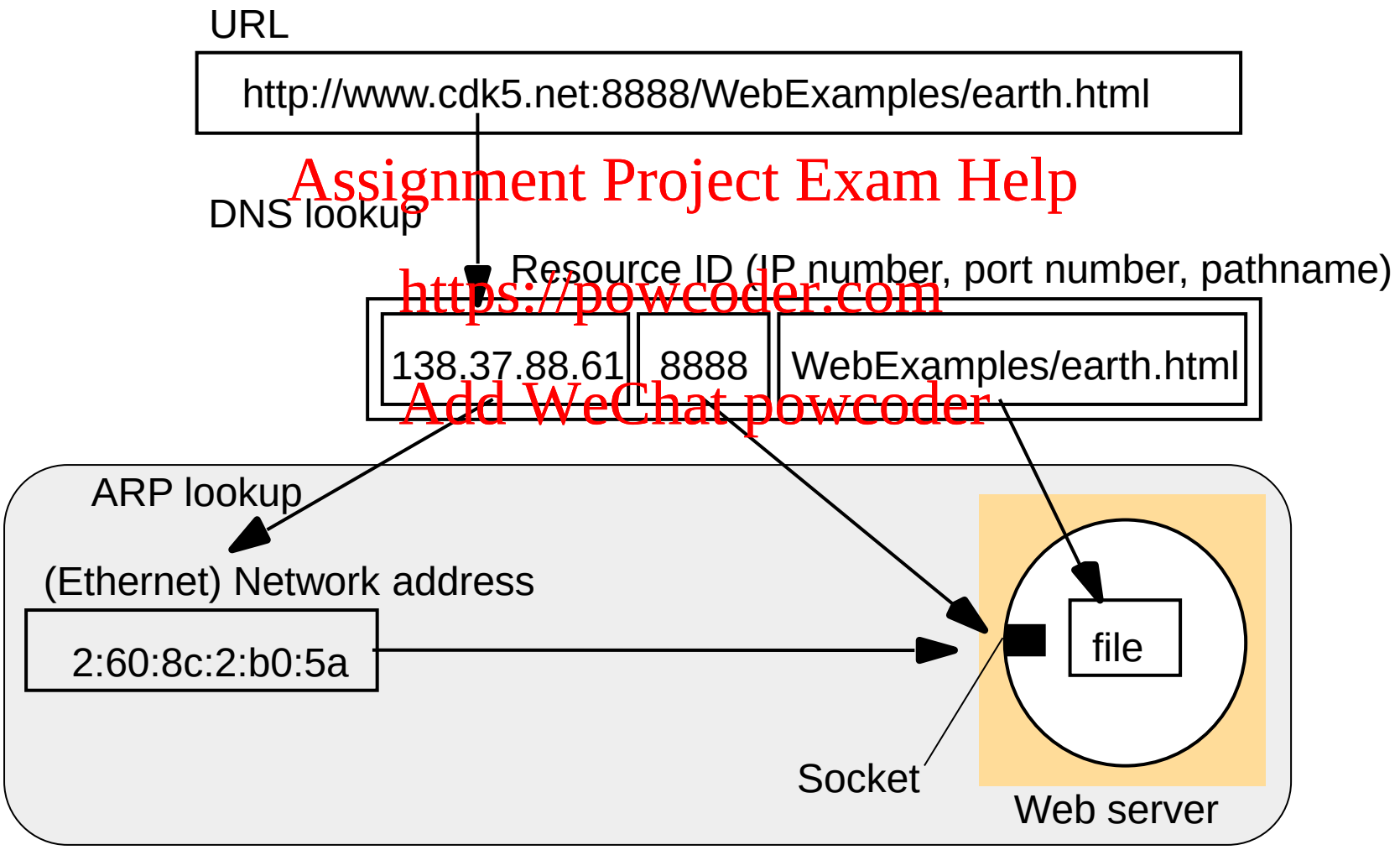
The Role of Names and Name Services

- Resources are accessed using *identifier* or *reference*
  - An identifier can be stored in variables and retrieved from tables quickly
  - Identifier includes or can be transformed to an address for an object
    - E.g. NFS file handle, CORBA remote object reference
  - A *name* is human-readable value (usually a string) that can be *resolved* to an identifier or address
    - Internet domain name, file pathname, process number
    - E.g ./etc/passwd, http://www.cdk5.net/
- For many purposes, names are preferable to identifiers
  - because the binding of the named resource to a physical location can be changed
  - because they are more meaningful to users
- Resource names are *resolved* by name services
  - to give identifiers and other useful attributes

Name Resolution



Accessing Resources from URL



Names and Resources

- Currently, different name systems are used for each type of resource:

resource	*name*	*identifies*
file	pathname	file within a given file system
process	process Id	process on a given computer
port	port number	IP port on a given computer

- Uniform Resource Identifiers (URI) offer a general solution for any type of resource. There two main classes:

URL Uniform Resource Locator (URL)

- typed by the protocol field (http, ftp, nfs, etc.)
- part of the name is service-specific
- resources cannot be moved between domains

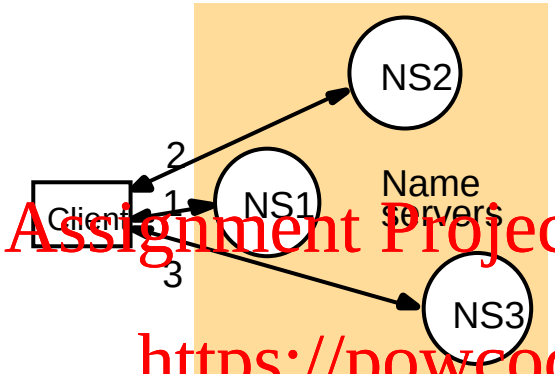
URN Uniform Resource Name (URN)

- requires a universal resource name lookup service - a DNS-like system for all resources

Navigation

- Navigation is the act of chaining multiple Naming Services in order to resolve a single name to the corresponding resource.
- Namespaces allows for structure in names.
- URLs provide a default structure that decompose the location of a resource in
  - protocol used for retrieval
  - Internet end point of the service exposing the resource
  - service specific path
- This decomposition facilitates the resolution of the name into the corresponding resource
- Moreover, structured namespaces allows for iterative navigation…

Iterative Navigation



A client iteratively contacts name servers NS1–NS3 in order to resolve a name

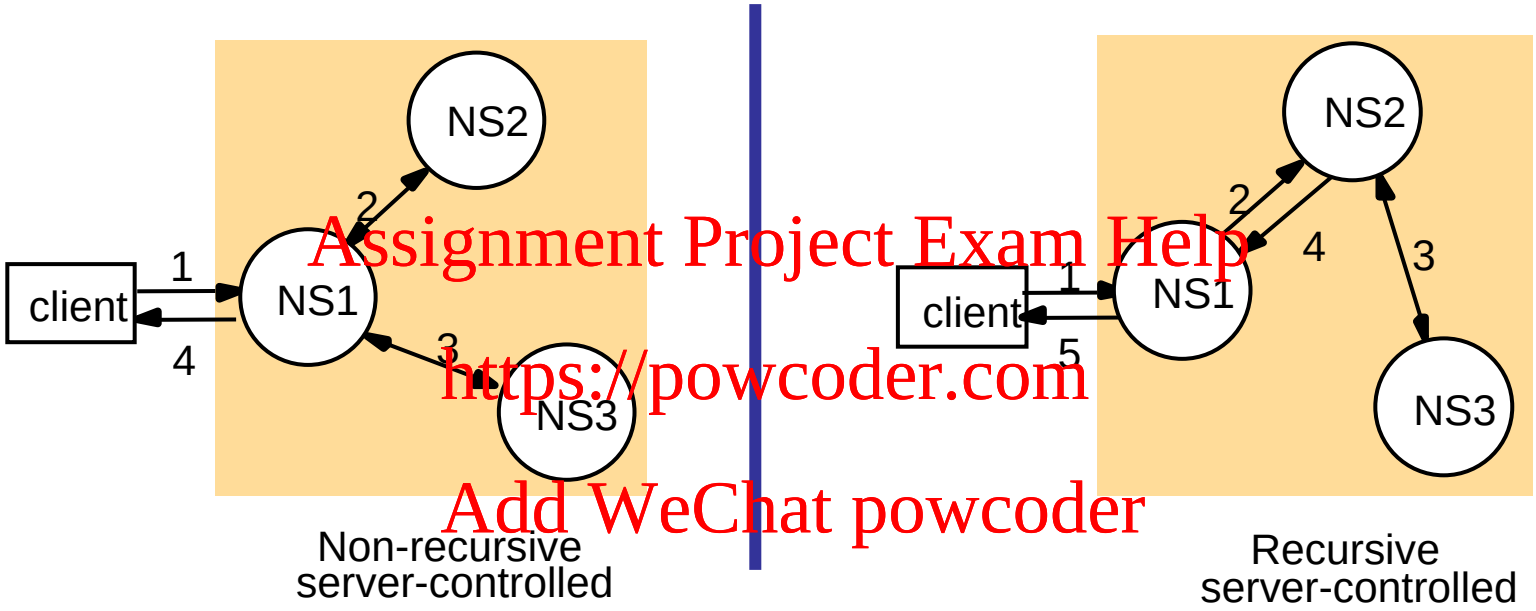
Used in:

- **DNS:** Client presents entire name to servers, starting at a local server, NS1. If NS1 has the requested name, it is resolved, else NS1 suggests contacting NS2 (a server for a domain that includes the requested name).
- **NFS:** Client segments pathnames (into 'simple names') and presents them one at a time to a server together with the filehandle of the directory that contains the simple name.

Server Controlled Navigation

- In an alternative model, name server coordinates naming resolution and returns the results to the client. It can be:
  - **Recursive:**
    - it is performed by the naming server
    - the server becomes like a client for the next server
    - this is necessary in case of client connectivity constraints
  - **Non recursive:**
    - it is performed by the client at the first server
    - the server bounces back the next hop to its client

Non-recursive and Recursive Server Controlled Navigation



A name server NS1 communicates with other name servers on behalf of a client

DNS offers recursive navigation as an option, but iterative is the standard technique. Recursive navigation must be used in domains that limit client access to their DNS information for security reasons.

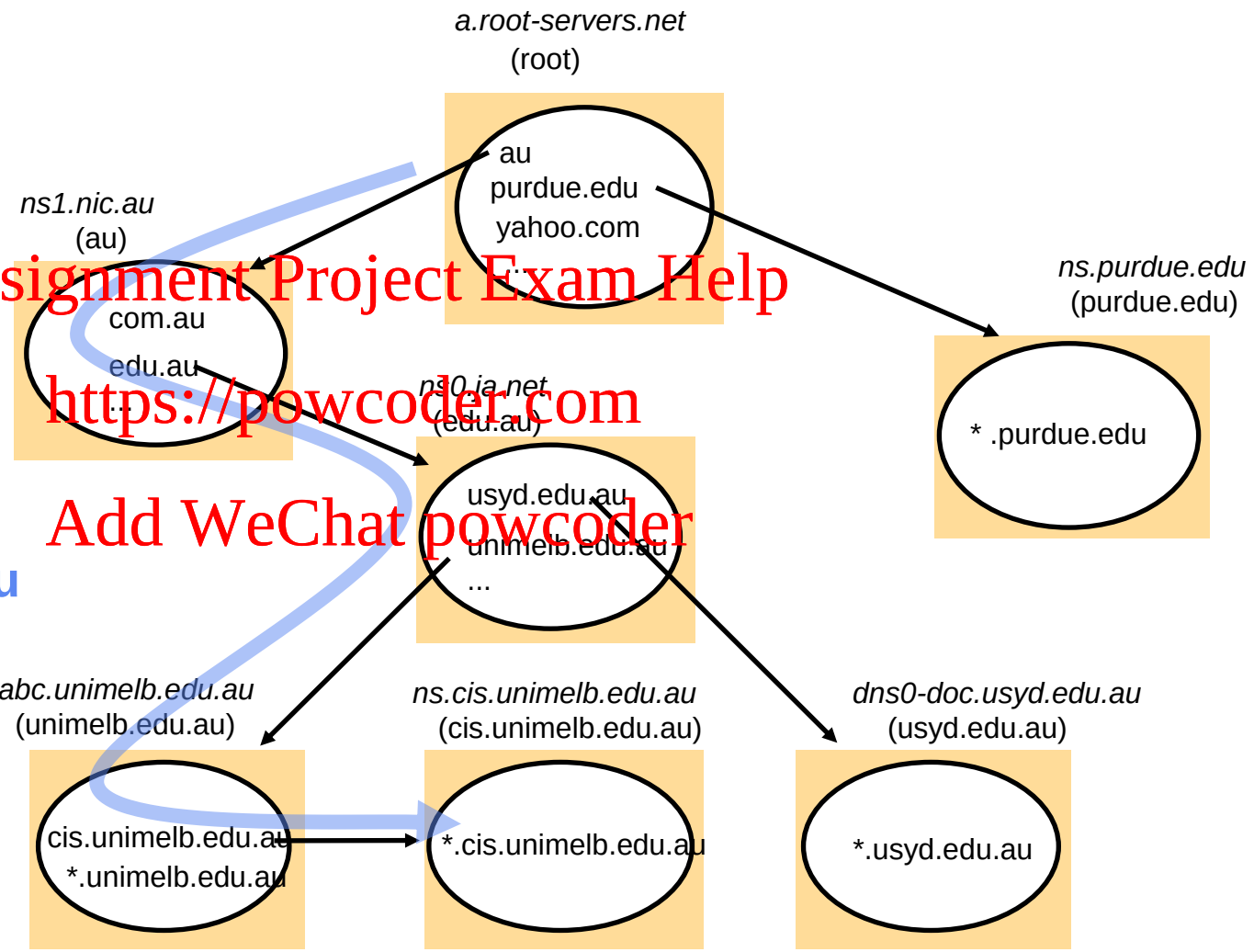
The Domain Name System (DNS)

- A distributed naming database (specified in RFC 1034/1035)
- Name structure reflects administrative structure of the Internet
- Rapidly resolves domain names to IP addresses
  - exploits caching heavily
  - typical query time ~100 milliseconds
- Scales to millions of computers: partitioned database, caching
- Resilient to failure of a server: replication
- Basic DNS algorithm for name resolution (domain name -> IP number):
  - Look for the name in the local cache
  - Try a superior DNS server, which responds with:
    - another recommended DNS server
    - the IP address (which may not be entirely up to date)

DNS Name Servers: Hierarchical Organisation

Note: Name server names are in italics, and the corresponding domains are in parentheses. Arrows denote name server entries

authoritative path to lookup:
dsys.cis.unimelb.edu.au



DNS Server Functions

- Main function is to resolve domain names for computers, i.e. to get their IP addresses
  - caches the results of previous searches until they pass their 'time to live'
- Other functions:
  - get *mail host* for a domain
  - reverse resolution - get domain name from IP address
  - Host information - type of hardware and OS
  - Well-known services - a list of well-known services offered by a host
  - Other attributes can be included (optional)

DNS Resource Records

Record type	*Meaning*	*Main contents*
A	A computer address (IPv4)	IPv4 number
AAA	A computer address (IPv6)	IPv6 number
NS	An authoritative name server	Domain name for server
CNAME	The canonical name for an alias	Domain name for alias
SOA	Marks the start of data for a zone	Parameters governing the zone
PTR	Domain name pointer (reverse lookups)	Domain name
HINFO	Host information	Machine architecture and operating system
MX	Mail exchange	List of <*preference, host*> pairs
TXT	Text string	Arbitrary text

DNS Issues

- Name tables change infrequently, but when they do, caching can result in the delivery of stale data.
  - Clients are responsible for detecting this and recovering
- Its design makes changes to the structure of the name space difficult. For example:
  - merging previously separate domain trees under a new root
  - moving subtrees to a different part of the structure (e.g. if Scotland became a separate country, its domains should all be moved to a new country-level domain.)

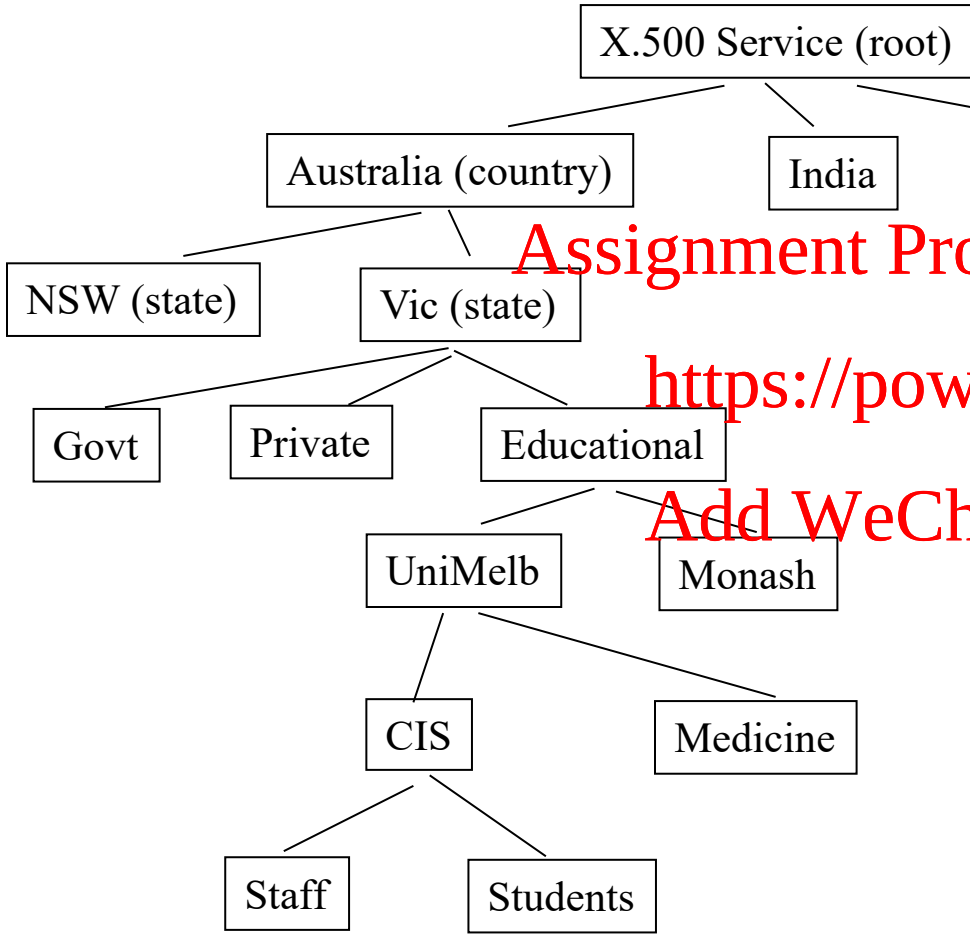
Directory Services

- Sometime users wish to find a particular person or resource, but they don't know its name, only some of its attributes.
  - What is the name of the user with a telephone number 03-83441344?
  - What is the name of an academic researching Cloud computing at UniMelb (e.g., ask Google!)
- Sometime users require a service, but they are not concerned with what system entity provides it.
  - Where can I print high resolution colour image?
- Directory services can help with above situation: they store collections of bindings and attributes and also looks up entries that match attribute-based specs.
- Directory service: 'yellow pages' for the resources in a network
  - Retrieves the set of names that satisfy a given description
  - e.g. X.500, LDAP, MS Active Directory Services
- Discovery service:- a directory service that also:
  - is automatically updated as the network configuration changes
  - discovers services required by a client (who may be mobile) within the current scope, for example, to find the most suitable printing service for image files after arriving at a hotel.
  - Examples of discovery services: Jini discovery service, the 'service location protocol'

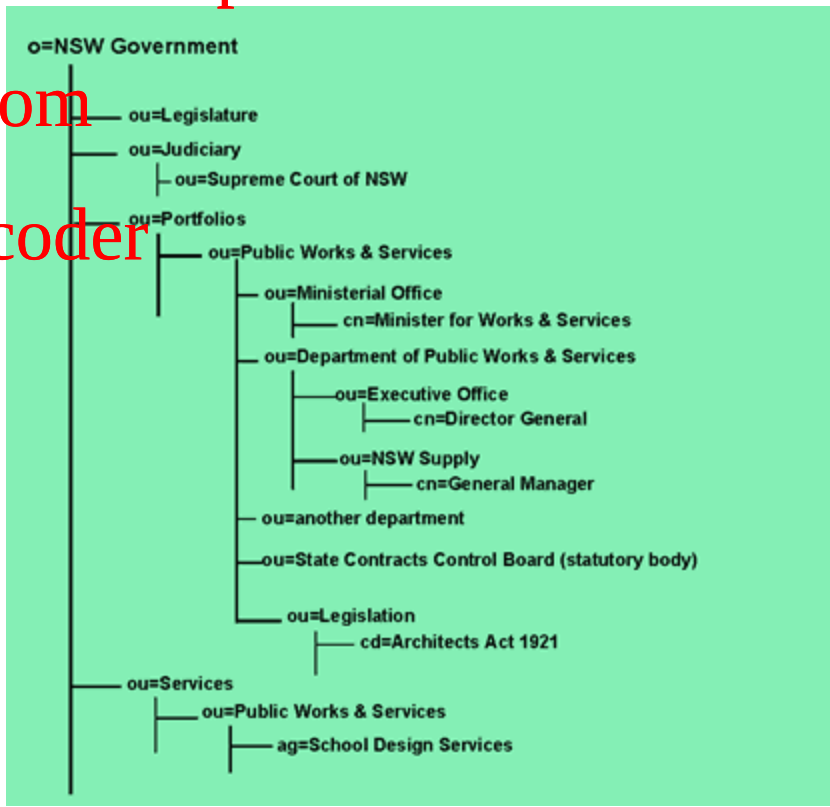
X.500 Directory Service

- X.500 and LDAP (Lightweight Directory Access Protocol)
  - a hierarchically-structured standard directory service designed for world-wide use
  - X.500 is standardised by ITU (International Telecommunication Union) and ISO
  - accommodates resource descriptions in a standard form and their retrieval for any resource (online or offline)
  - never fully deployed, but the standard forms the basis for LDAP, the Lightweight Directory Access Protocol, which is widely used – IETF RFC 2251.
  - A secure access to directory through authentication is also supported.

Part of the X.500 Directory Information Tree (DIT)



Object class for NSW govt.



Summary

- Names services facilitate communication and resource sharing in distributed systems.
- They are playing an important role in Distributed systems such as the Internet, Web, CDNs (Content Delivery Networks), Web Services, Location-aware services – publication and discovery
- Name services:
  - defer the binding of resource names to addresses (and other attributes)
  - Names are resolved to give addresses and other attributes
  - Goals :
    - Scalability (size of database, access traffic (hits/second), update traffic)
    - Reliability
    - Trust management (authority of servers)
    - exploitation of replication and caching to achieve scalability without compromising the distribution of updates
- Directory and discovery services:
  - 'yellow pages' retrieval by attributes
  - dynamic resource registration and discovery