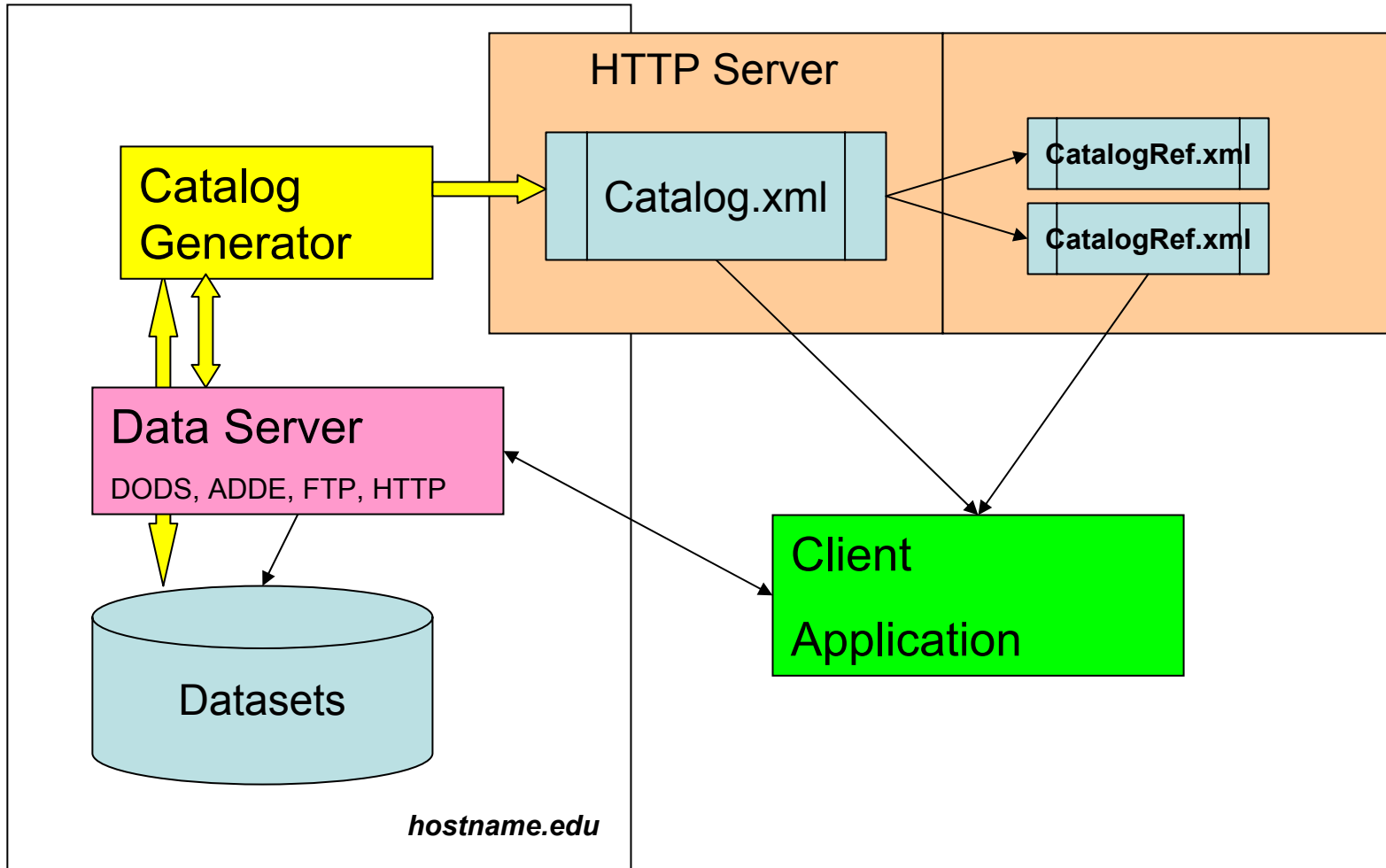


Recent Work in Progress

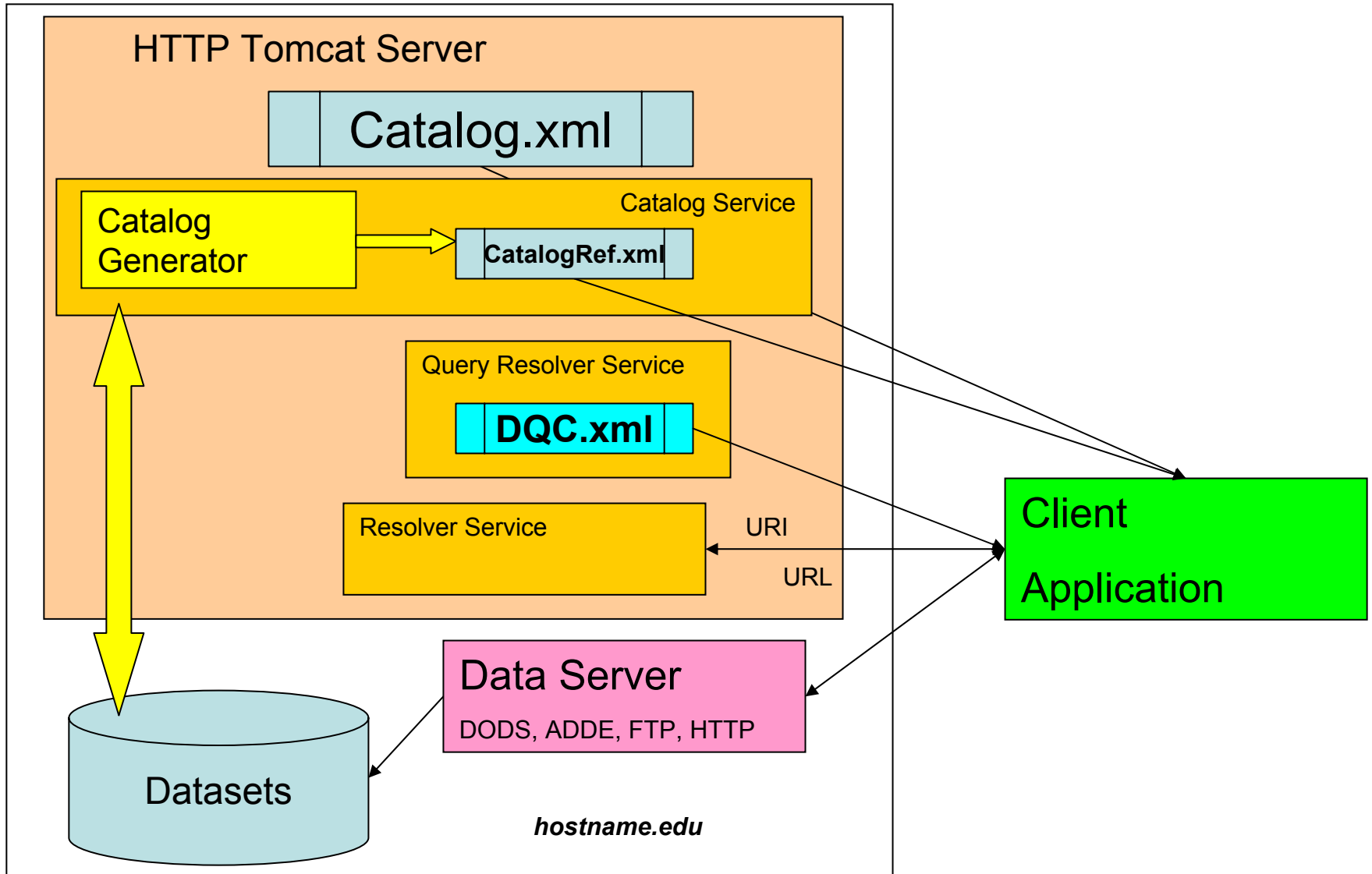
John Caron, June 3, 2003

- THREDDS development
 - Dynamic Catalogs: DQC, Resolvers
 - IDD Data Server
 - ADDE Cataloger
- NetCDF development
 - NetCDF Markup Language (NcML)
 - More efficient Java I/O (NIO)
 - NetCDF/DODS/HDF5 Data Models

THREDDS Catalogs



Dynamic Catalogs = Services



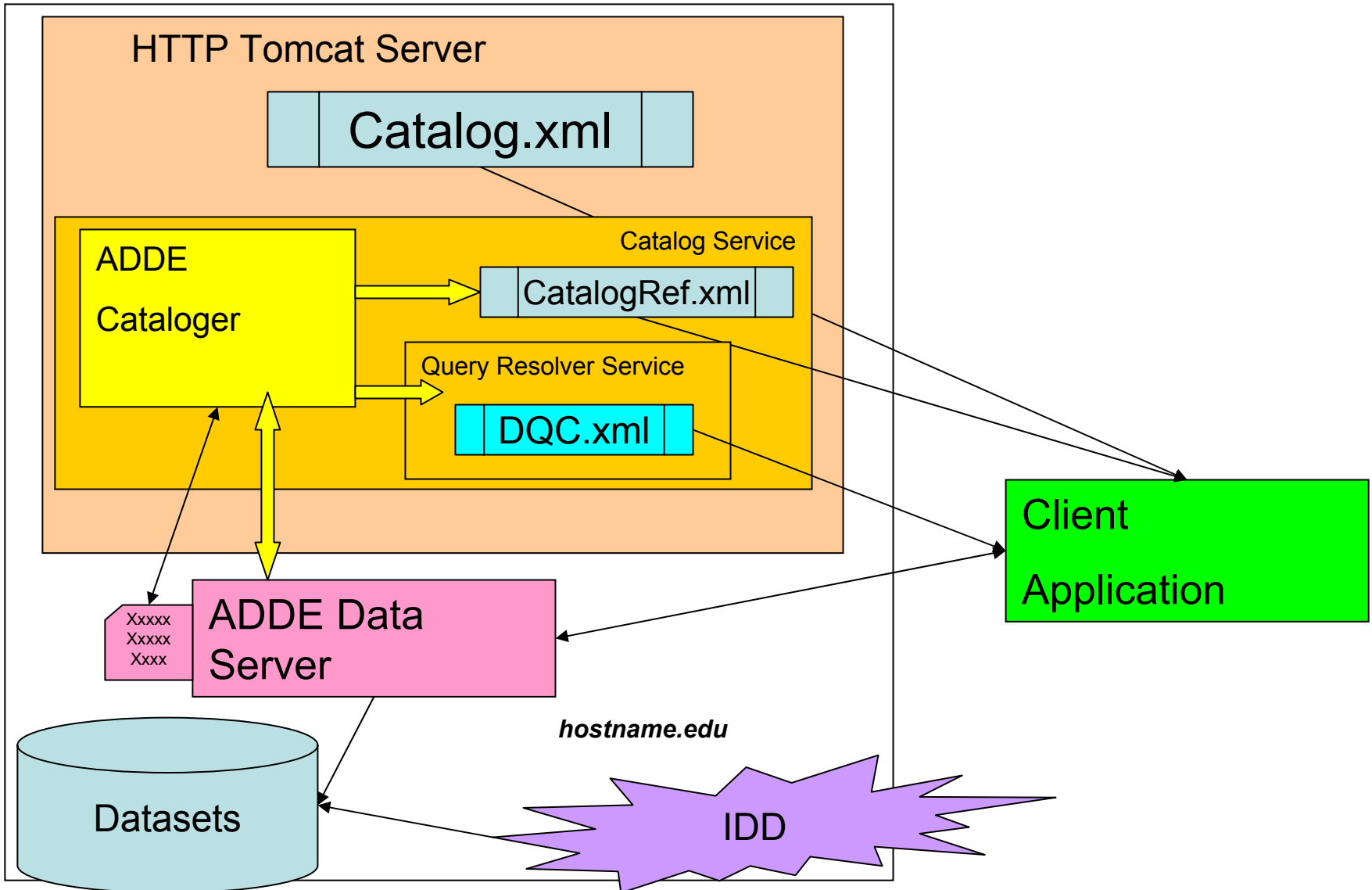
Dataset Query Capability (DQC)

- XML document.
- Describes what the user can ask for as a set of orthogonal “selections”.
- On the client, a “query URL” is formed based on the user’s choices, and sent to the server.
- The “query resolver” server finds which datasets satisfy the query and returns a list of real dataset URLs.
- The DQC describes the queries that the server is capable of responding to.

Resolver Services

- Logical Dataset, eg “latest ETA model run”
- Dataset with Service type “Resolver”
- On the client, the URI of the logical dataset is sent to the server
- The server finds what is available and returns a list of real dataset URLs.

ADDE Cataloger

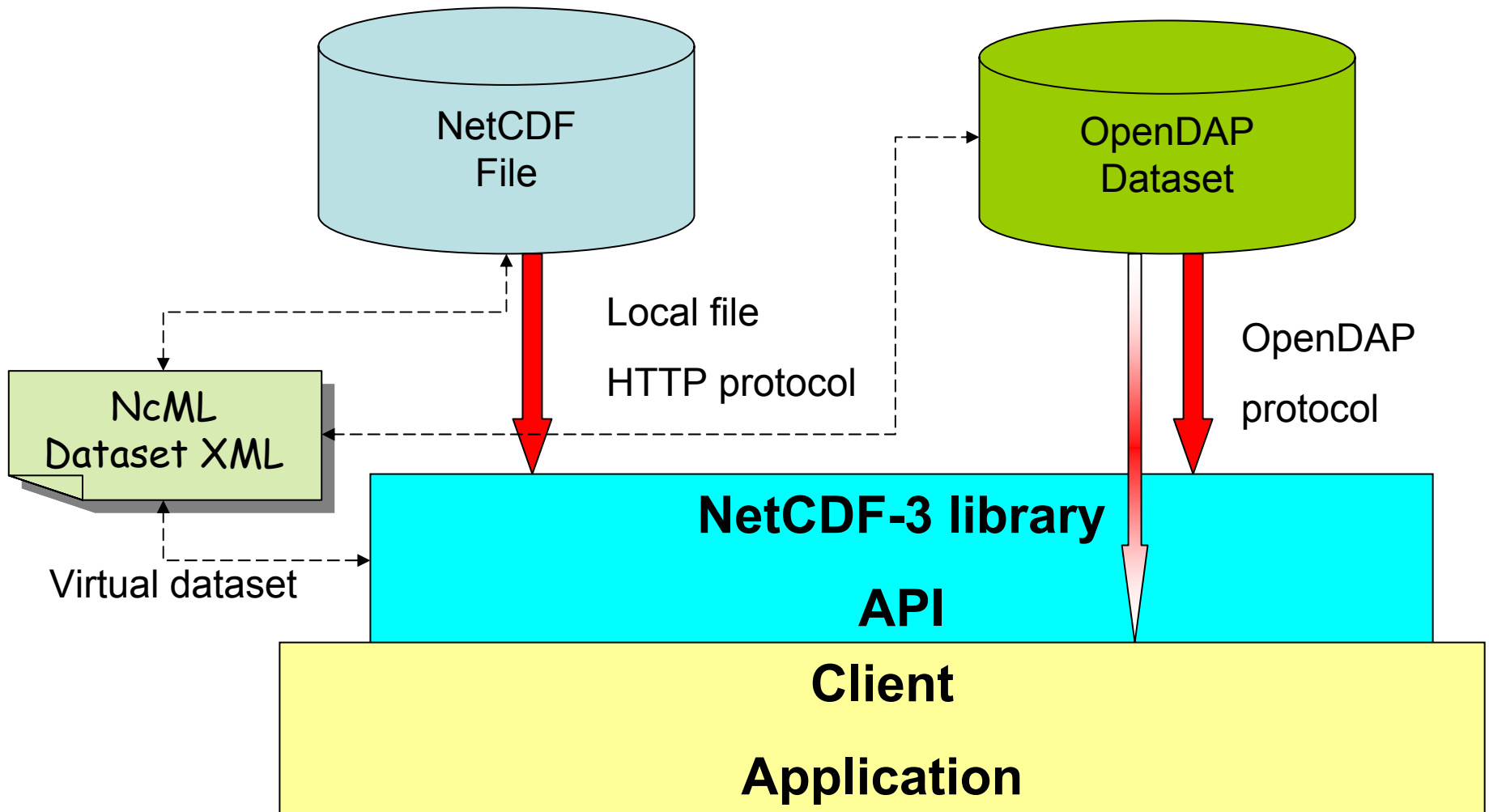


Summary IDD Data Server

Get as much of the IDD Data feeds available via THREDDS as possible.

- NCEP model data (catgen) (DODS)
- Level 3 NEXRAD (custom server/DQC) (ADDE)
- SSEC/Unidata Satellite data (ADDE Cataloger) (ADDE)
- Text Data: Metars, Surface Obs, etc (DQC/custom server), returns text or XML.
- Profiler Data (custom server/DQC) (ADDE)

NetCDF 3

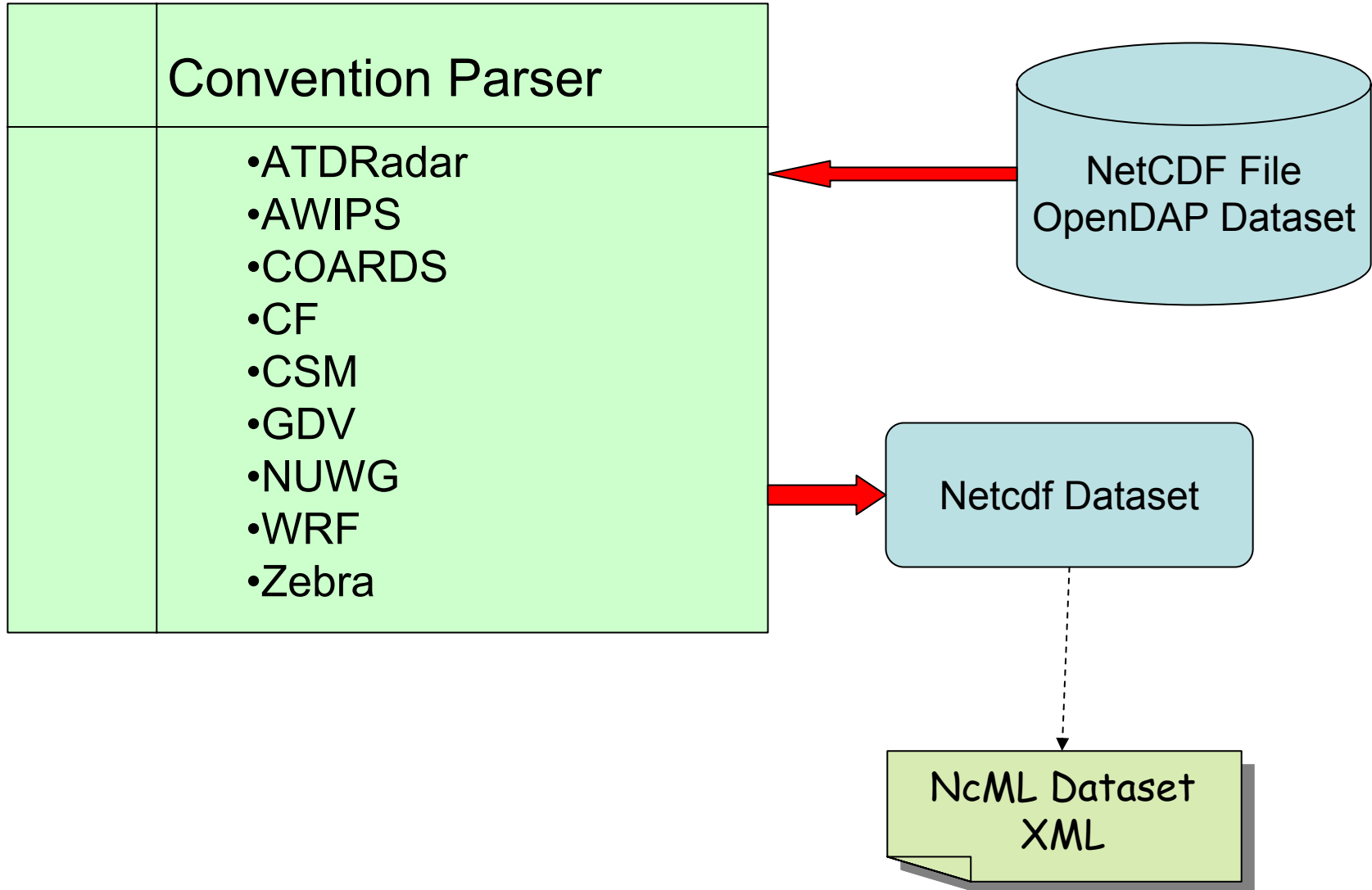


NetCDF Markup Language

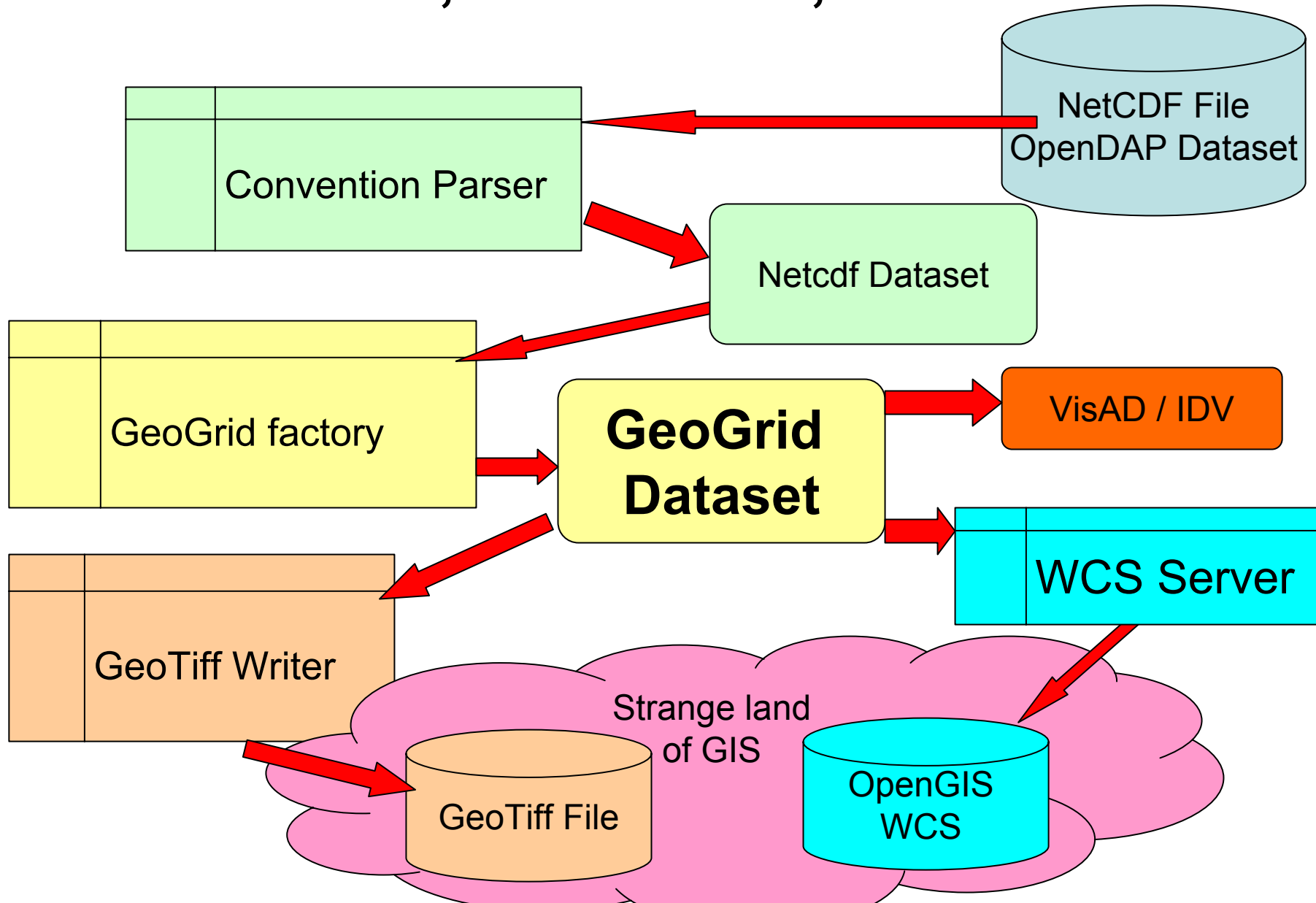
XML representation of netCDF metadata, uses XML Schema

- **Core:** existing netCDF data model
- **Coordinate System:** general and georeferencing coordinate system
- **Dataset:** redefine, aggregate, subset
- *Luca Cinquini (NCAR/SCD/ESG), John Caron, Ethan Davis, Bob Drach (LLNL), Stefano Nativi (Florence), Russ Rew*

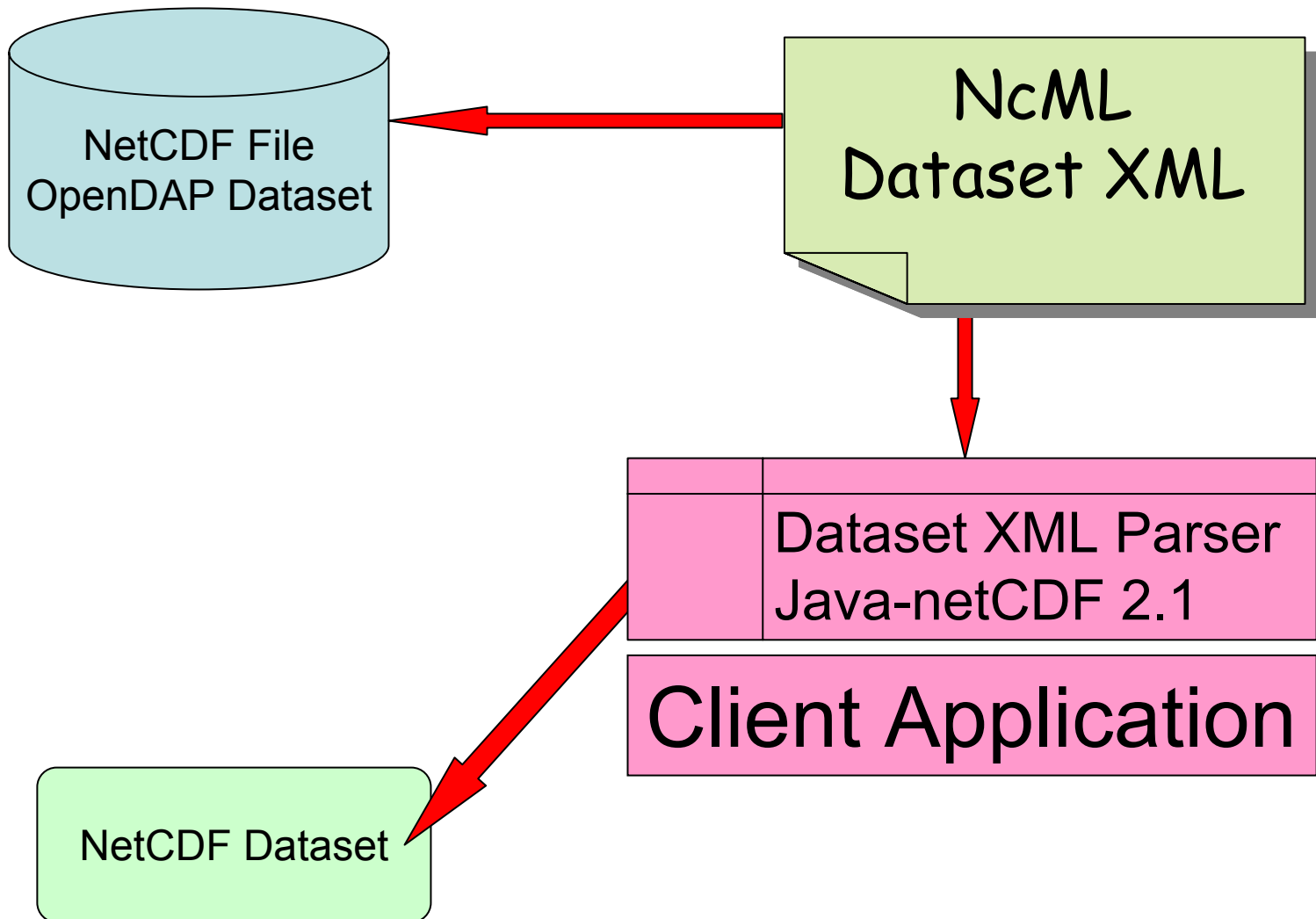
NcML Coordinate Systems



GeoGrids, GeoTiffs, Geowhiz!



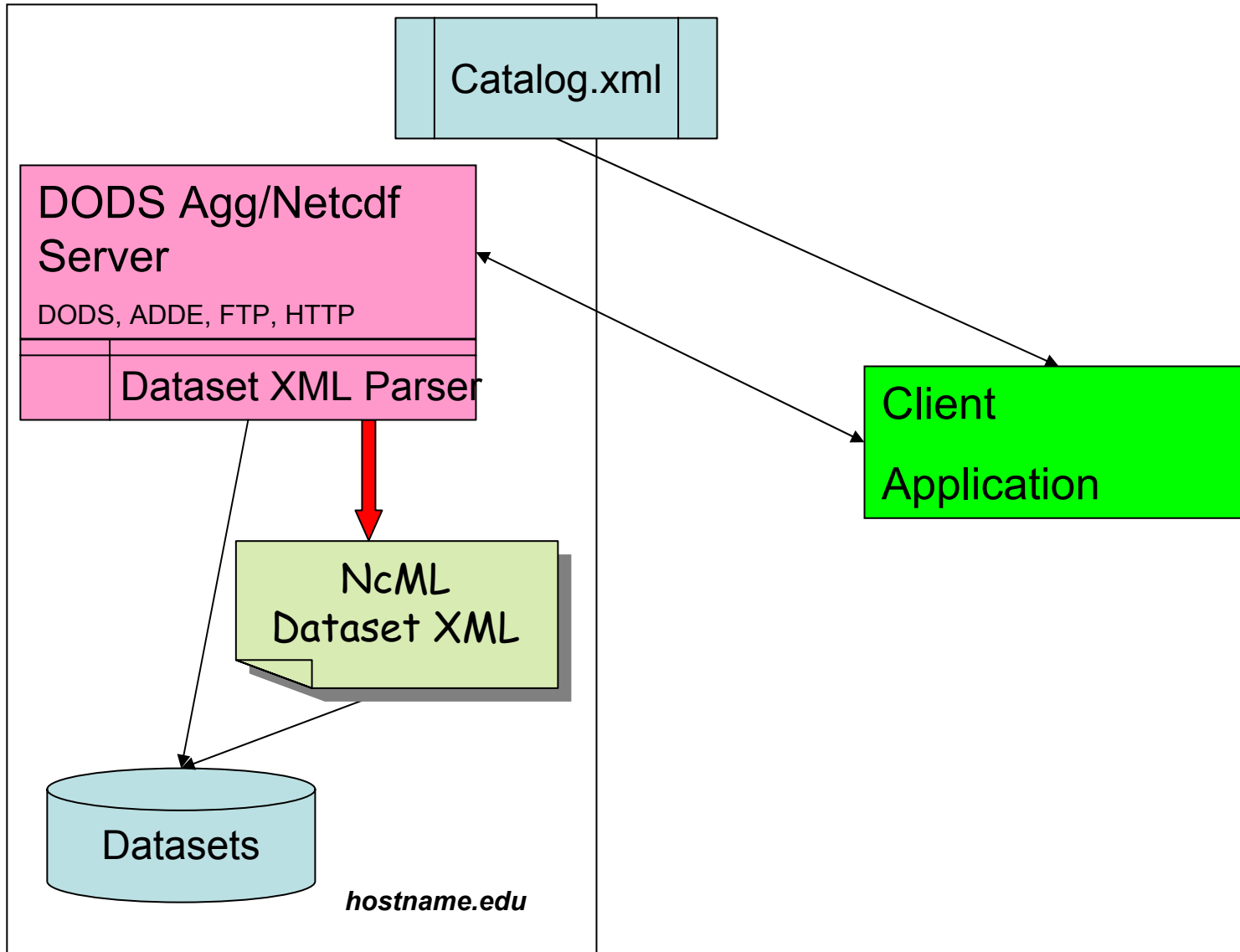
NcML Dataset : “virtual view”



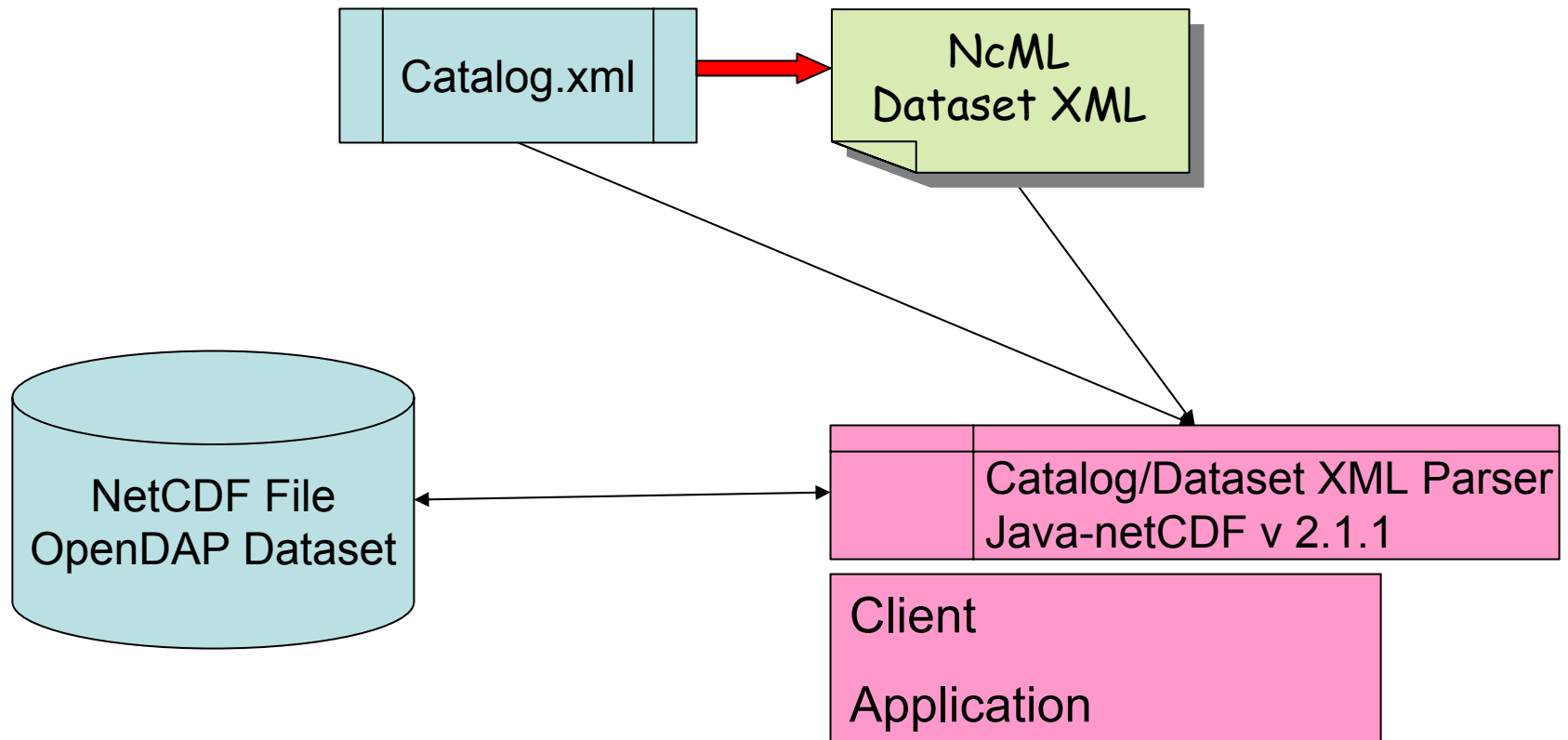
NcML Dataset

- Use NcML like CDL, to declare the contents of a netCDF file.
- Add, delete or rename Variables, Attributes, and Dimensions
- Subset Variables
- Reorder a Variable's dimensions
- Aggregate multiple netCDF files, a la DODS Aggregation Server
- NcML Dataset is a “virtual view” or can make copy to a local netCDF file.

2: NcML Datasets on a Server



3: NcML Datasets via Catalogs



NIO

- Rewrite ucar.nc2 I/O layer using java.nio package (currently using ucar.netcdf)
- Uses memory mapping, bulk I/O transfer
- Prototype has 7x speedup on large files.
- Requires JDK 1.4+
- HTTP access must be rewritten

NIO vs current Java

	NIO	Current	old/new
First access			
small (3.9 Mb)	281	671	2.4
large (240 Mb)	3334	28221	8.5
Average next 5 accesses			
small	54	290	5.4
large	2239	16367	7.3

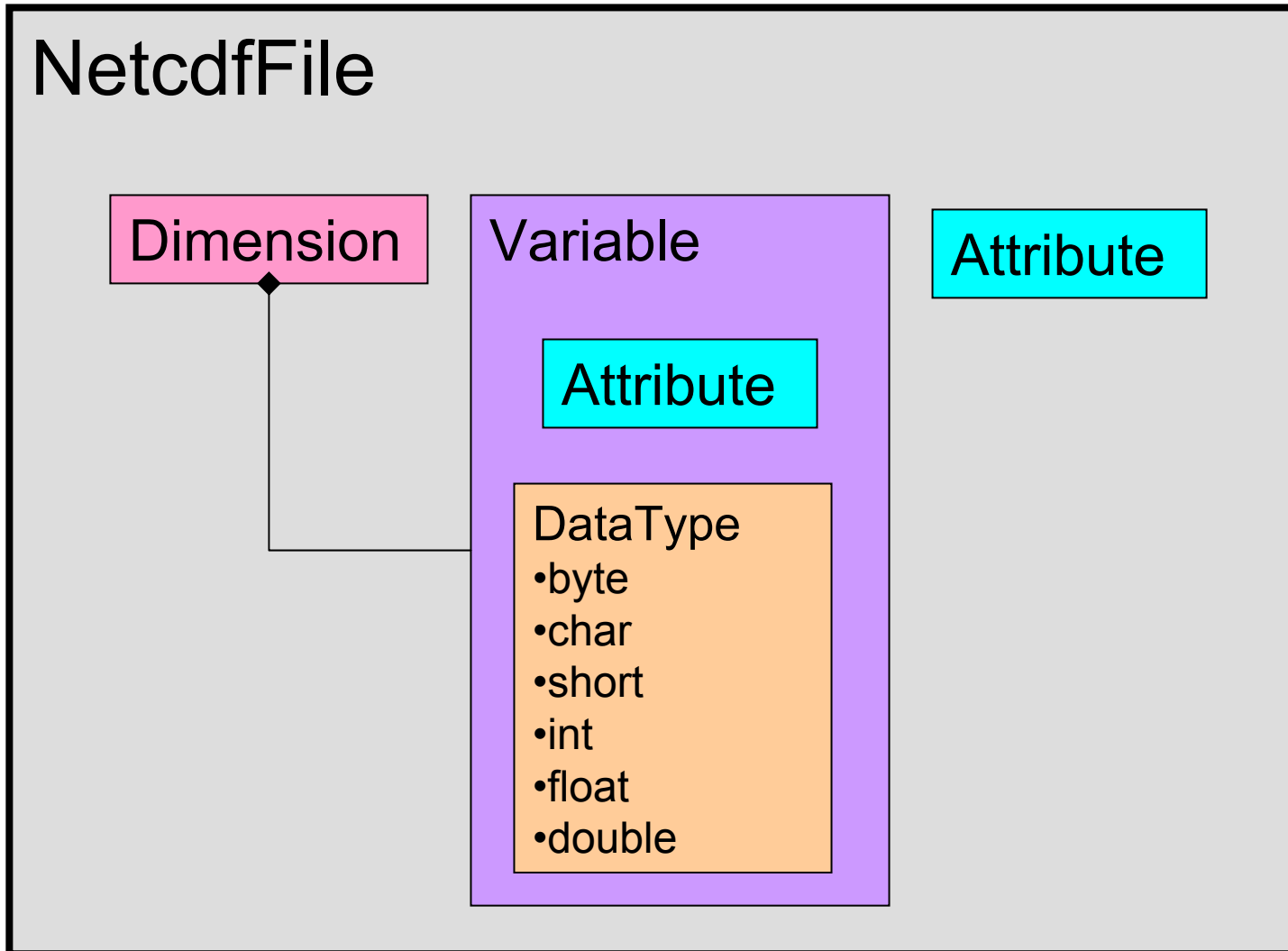
- Time in millisecs to sequentially read entire file
- Wintel 2GHz, 1 GB main memory
- Java 1.4.2 -client

NIO vs optimized C

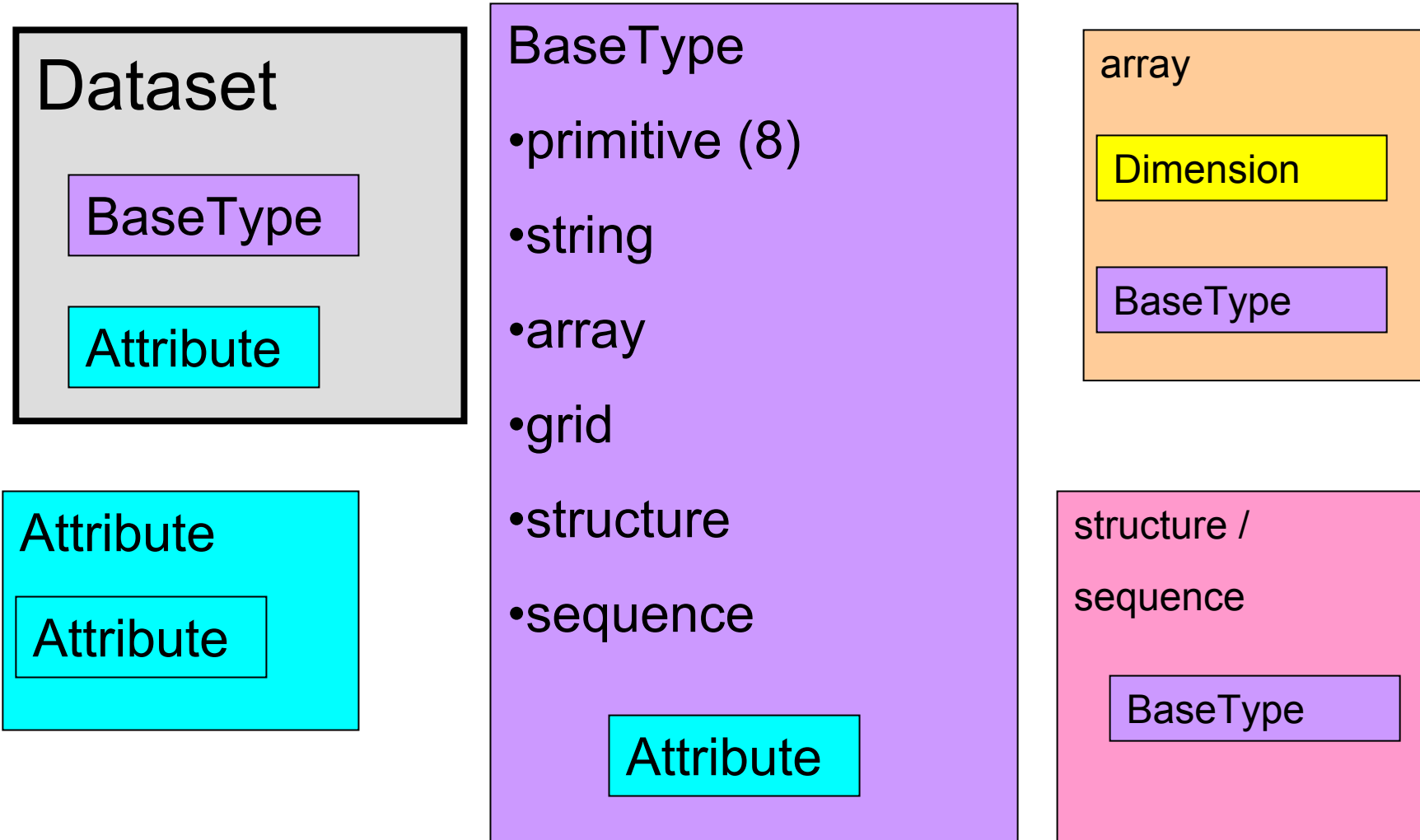
	NIO	C	C/NIO
First access			
small (3.9 Mb)	281	370	1.3
large (240 Mb)	3334	19348	5.8
Average next 5 accesses			
small	54	24	.44
large	2239	953	.43

- Java 1.4.2 –client vs. VC 6.0 /O2

NetCDF Data Model



OpenDAP Data Model



HDF5 Data Model

Groups

File directory structure inside HDF file.

Datatype

- Fixed point
- floating point
- date/time
- string
- bit field
- Opaque
- Compound
- Reference
- Enumeration
- Variable length
- Array

Dataset

Datatype

DataSpace

Attribute

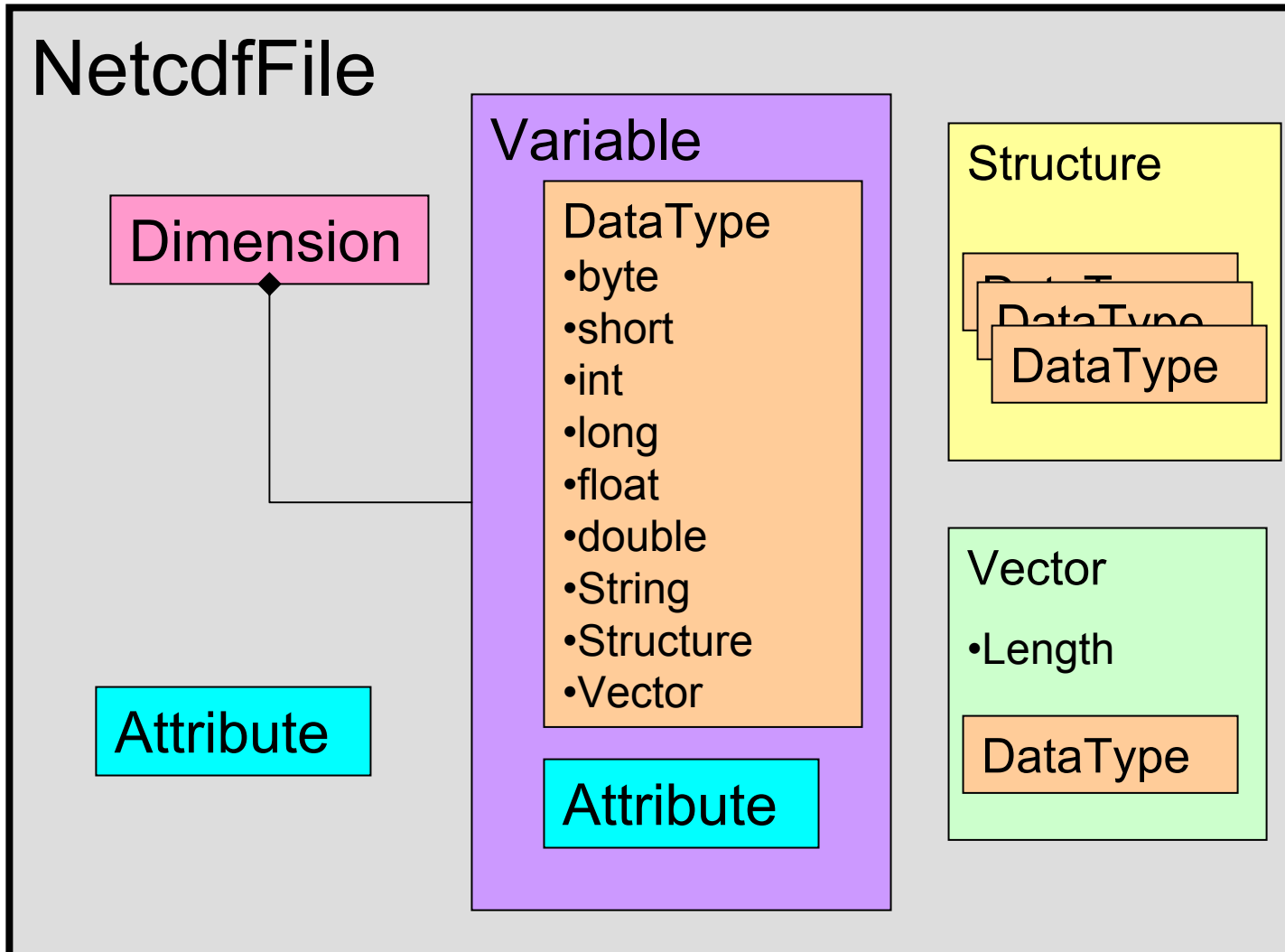
Data storage

- Compact
- External
- Layout
- Indexed
- Striped

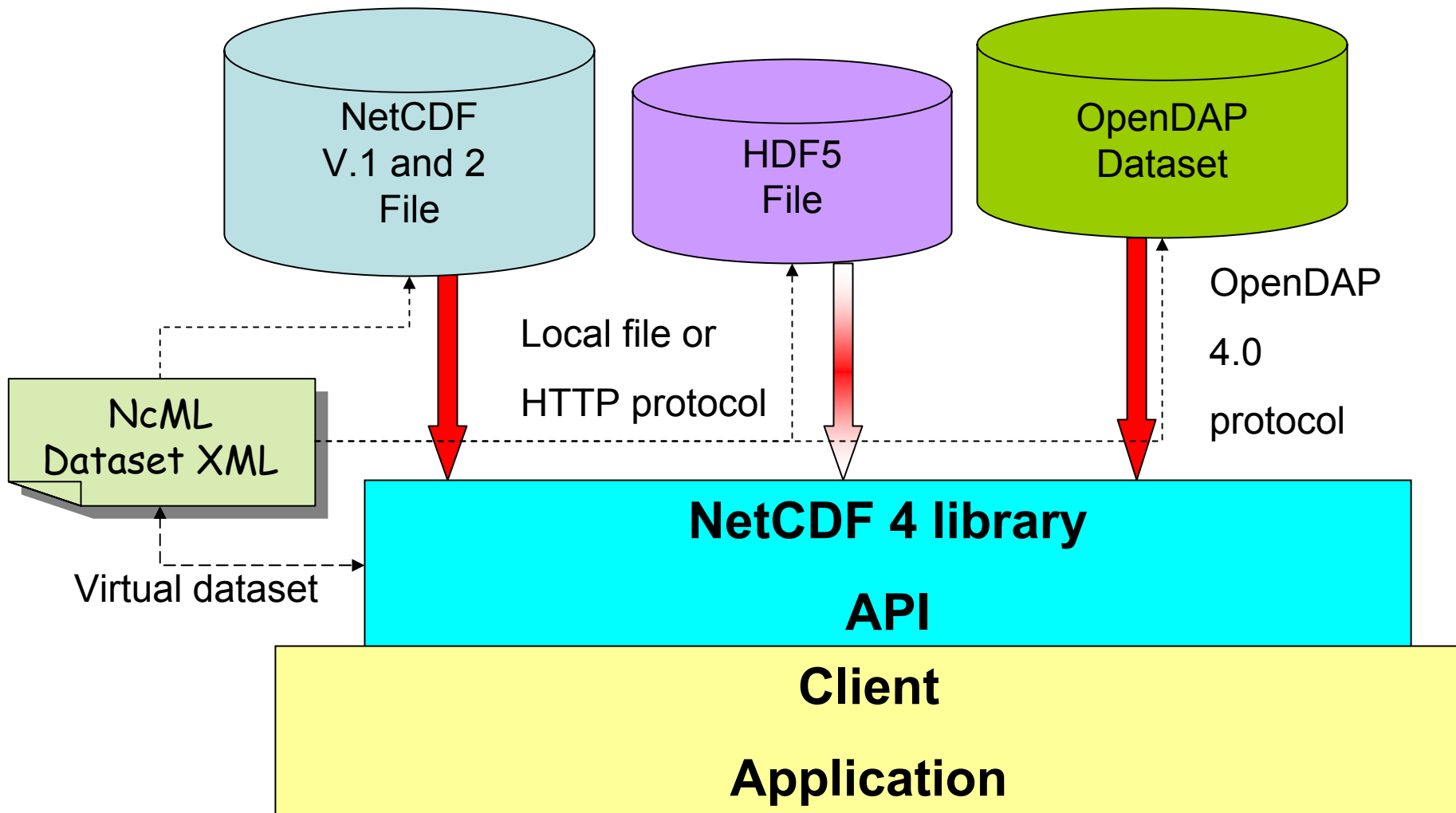
Possible Extensions to netCDF data model

- Add new data types:
 - **Strings**: variable length arrays of bytes, plus an encoding attribute.
 - **Structures**: collections of any other element types, allow nested structures.
 - **Vector**: a variable length 1D array of any type.
- Allow reusable structure definition = **user defined data type**.
- Allow unnamed, undeclared dimensions = **anonymous dimensions**.
- Allow multiple **unlimited dimensions** (outer dimension only)
- **Compression**. Push scale/offset into library, allow variable bit sizes.
- Explicit support for **coordinate variables/axes**.

New NetCDF Data Model



NetCDF 4





netCDF

One interface to bind them