Accomplishments and Remaining Challenges:

THREDDS Data Server and Common Data Model

Ethan Davis

Unidata Policy Committee Meeting

23-24 May 2011

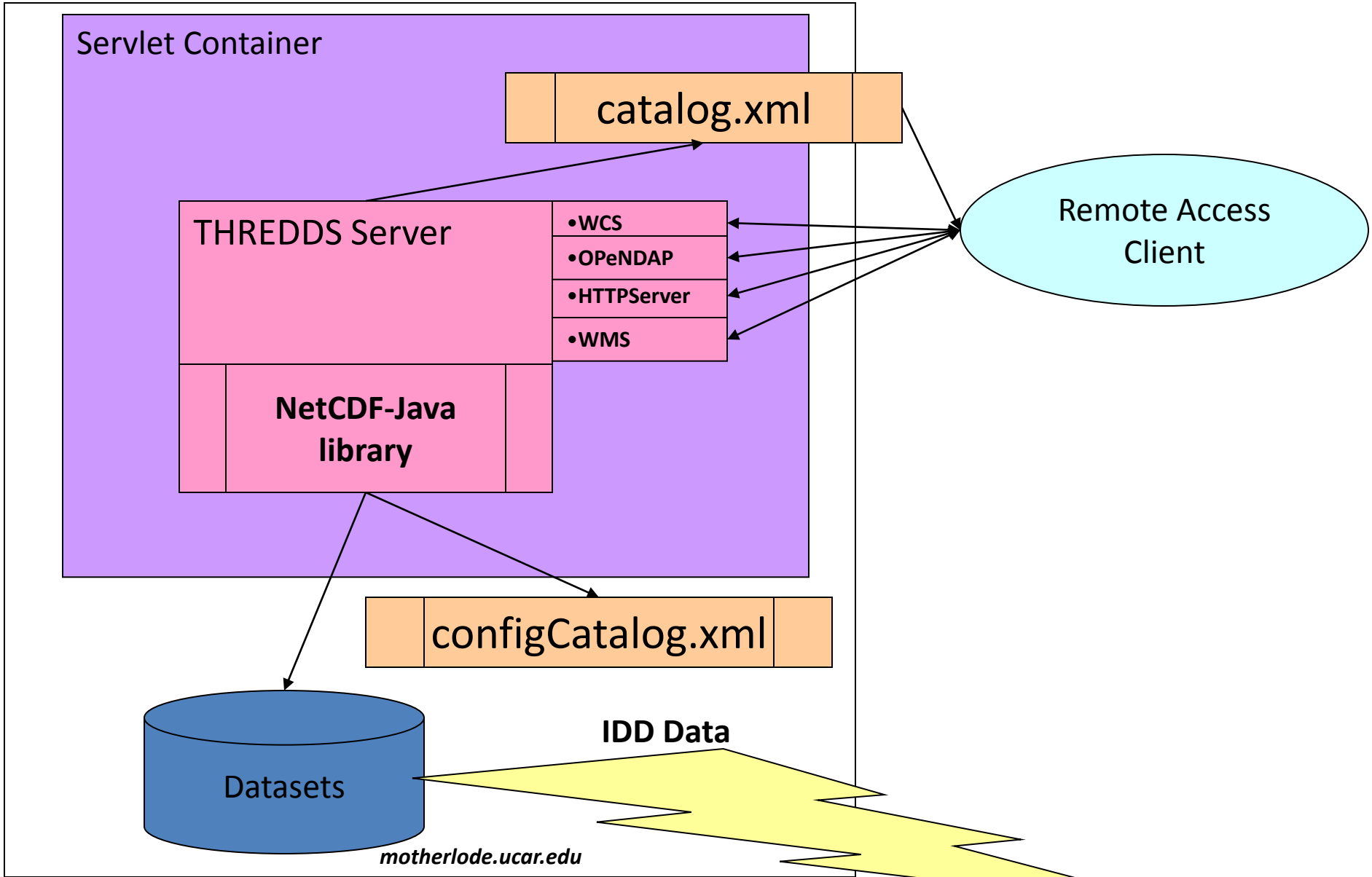
THREDDS Data Server (TDS)

- Web server for scientific data (written in 100% Java)
- Can serve any dataset the netCDF-Java library can read
  - E.g., netCDF-3, netCDF-4, HDF-4, HDF-5, HDF-EOS, GRIB-1, GRIB-2
- Catalogs of available datasets:
  - Hierarchical presentation of dataset collections
    - Human readable (HTML) representation
    - Machine readable (XML) representation
  - Advertises services available for datasets
  - Contains metadata about each dataset
- …

THREDDS Data Server (TDS)

- …
- Data access services:
  - OPeNDAP
  - OGC WMS and WCS
  - NCSS
- Data collection services
  - Aggregation
  - Point/station collection
- Metadata services
  - THREDDS
  - ncISO: ISO, UDDC, NcML

THREDDS Data Server

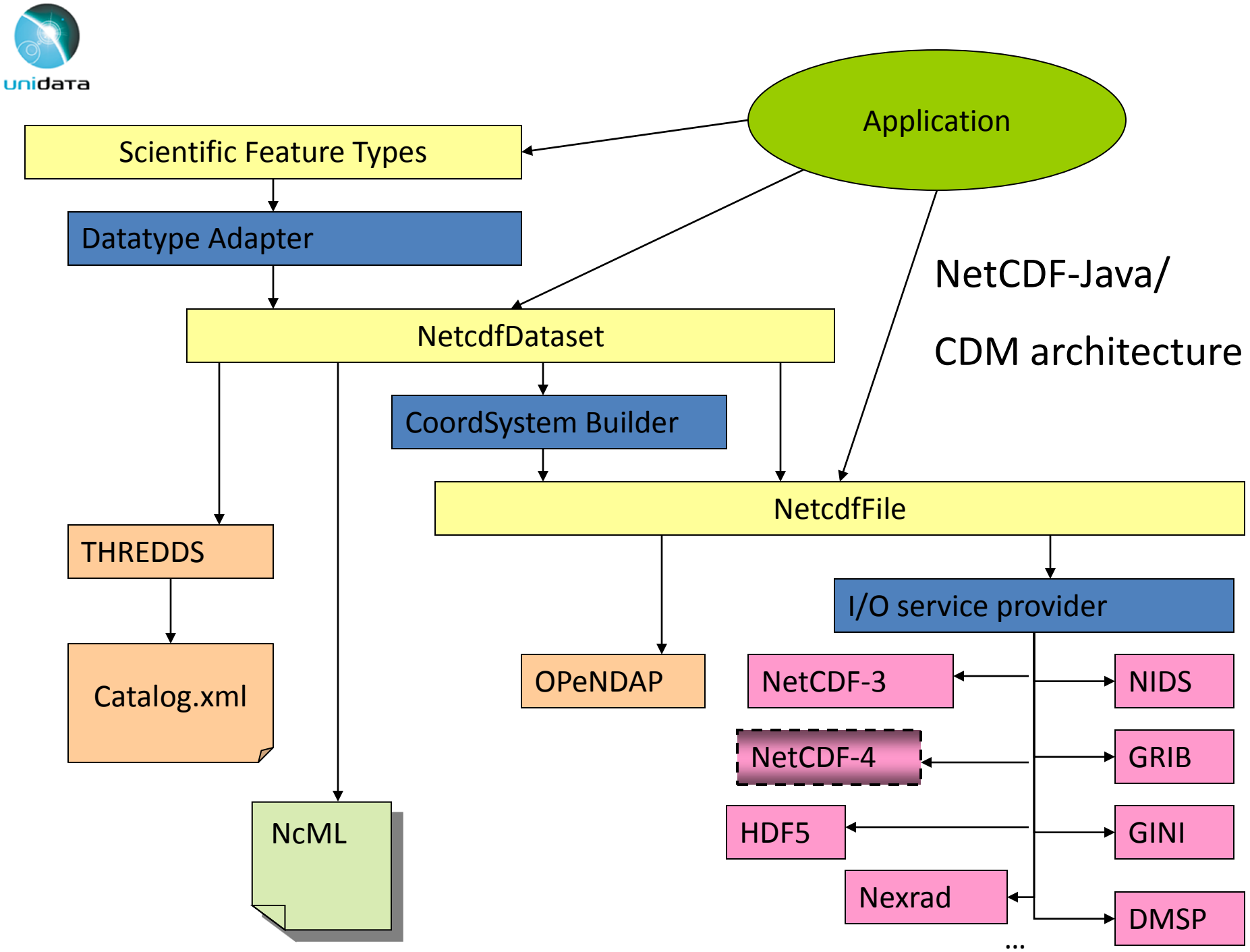


Common Data Model (CDM)

- Implemented in the netCDF-Java library (3 layers)
- NetCDF Data Access:
  - Access to array-based scientific data (netCDF data model)
  - Read various file formats, map into data model
- NetcdfDataset:
  - Geospatial coordinate systems (conventions)
  - NcML: Modify existing dataset
  - Aggregation of datasets
- Scientific Feature Types, e.g.:
  - Gridded Data [*]; Radial Data; Swath Data
  - Discrete Sampling Features:
    - Point Data; Station Data; Profile Data
    - Trajectory (e.g., aircraft track) Data

---

[*] both “regular” and unstructured grids


Application
Scientific Feature Types
Datatype Adapter
NetcdfDataset
CoordSystem Builder
NetcdfFile
NetCDF-Java/
CDM architecture
THREDDS
Catalog.xml
NcML
OPeNDAP
I/O service provider
NetCDF-3
NetCDF-4
HDF5
Nexrad
...
NIDS
GRIB
GINI
DMSP

General proposal objectives related to CDM/TDS

- Provide … simple mechanisms for locating, accessing, and distributing real-time and thematic data, creating and publishing metadata
  - CDM Feature Types, THREDDS metadata, ncISO, GI-Cat (OGC CS/W)
- …
- Empower community members to create and deploy innovative data services
  - TDS, netCDF, CF conventions, OPeNDAP, OGC WMS, NCSS
- …

Specific CDM/TDS objectives: mostly completed

- Provide full data access to the HDF family of file formats (HDF/HDF-EOS, HDF5/HDF5-EOS/NPOESS)
- Complete implementation of scientific feature types (grid, point, station, track, radial, swath, unstructured grids) and standardized conventions to represent these feature types
- Make the TDS efficiently scale to large collections of datasets and files
- Serve weather and climate data to other communities:
  - Support OGC and ISO standards (WCS, WMS, WFS) to support users in other disciplines, e.g., GIS, data assimilation, numerical modeling
  - Enable users in other communities to access weather and climate data through their own client applications using standards-based protocols
- Map CF conventions to international standards specification such as the ISO “coverage”

Specific CDM/TDS objectives: partly completed

- Provide efficient access to output from ensemble forecast models
- Make Unidata/LDM data feeds and case studies available through the TDS and accessible by the IDV
- Create web services to subset large collections of data, returning data in convenient formats [use scientific feature types to guide aggregation and collection services]
- Work with community to develop data discovery metadata standards
- Improve automated metadata generation and publication
- Make progress, through appropriate collaborations, towards “Scientific Databases” to allow complex search capabilities over very large collections of scientific datasets.
- Facilitate the installation of the TDS in universities so that faculty and students can share their local data holdings

Specific CDM/TDS objectives: not yet started

- Provide efficient access to output on unstructured grids
- Enhance the CDM/TDS to provide support for aggregating, accessing as a collection, fusing and analyzing data from ensemble model output
- Providing data availability notifications

Specific CDM/TDS tasks: not anticipated in proposal

- Security enhancements suggested by NOAA's security audit of the TDS
- Enhance CDM to handle GRIB time interval variables (e.g., accumulated precipitation)
- Dataset access control in TDS (e.g., support ESG-style access control)
- Improve BUFR handling code in attempt to handle complexity/problems with BUFR data

Conclusion

- Some headway made on all objectives listed in proposal
- Many proposed and new objectives complex, experimental, high risk
- Limited staffing/resources, must prioritize
- Must also enable/entrain collaborators to accomplish more objectives