Unidata TDS Workshop

THREDDS Data Server Overview

23-24 July 2015
Unidata: Core Activities

• Facilitate access to (real-time) data by the University research and education community
• Support the community in their use of the data
• Help build, represent, and advocate on behalf of the community
Unidata: Facilitate access to real-time data
Unidata: Facilitate access to real-time data

Push

IDD: Real-Time Data Distribution

Over 200 sites. Approx 15 GB/hour

Unidata’s LDM

- Protocol and client/server software
- Event-driven data distribution
- Supports subscription to subsets of data feeds
Unidata: Facilitate access to real-time data
Pull

• IDD data from Unidata available via
  – Servers:
    • McIDAS ADDE
    • TDS
    • RAMADDA
  – Protocols:
    • HTTP, FTP
    • ADDE, OPeNDAP
    • OGC WCS and WMS

• The Unidata TDS server thredds.ucar.edu
  “archives” latest 30 days or so of IDD data
Unidata: Core Activities

• Facilitate access to (real-time) data by the University research and education community
• Support the community in their use of the data
• Help build, represent, and advocate on behalf of the community
Unidata: Core Activities

• Facilitate access to (real-time) data by the University research and education community
• Support the community in their use of the data
• Help build, represent, and advocate on behalf of the community

• Develop open source tools and infrastructure for data access, analysis, visualization, and data management
• Advance metadata standards for the earth science community
• Support users of our technologies
Visualizing and Analyzing Data

GEMPAK

McIDAS-X

IDV
Unidata User Community

• Support the community
  – User Workshops
  – Training Workshops
  – Mailing lists
    • For specific software packages
    • “community” email list – for Unidata community announcements

• Represent and advocate for the community

• More: http://www.unidata.ucar.edu/
Why TDS?

Users

• As a user, some things I want to be able to do:
  • Easily ‘see’ information regarding the dataset, without the need to download any files
    • Temporal / spatial ranges, available variables, contact info, dataset details
  • Get only the data I need/want
    • Temporal, spatial, and variable subsetting
  • Get data remotely in a variety of ways
  • Download one file, even if data span multiple files
Why TDS?

Users

* From PacIOOS site, developed by John Maurer, U of HI
Why TDS?

Users

* From PacIOOS site, developed by John Maurer, U of HI
Why TDS?
Data Providers

• As a data provider, I want to be able to:
  • Catalog my data holdings
  • Aggregate data files
  • Provide a ‘quick view’ of my data
  • Easily add information (metadata) to my datasets
  • Fix ‘incorrect’ datasets*
  • Allow flexibility in the way users access my data

Give users what they need to do science!
Why TDS?
Data Providers

- As a data provider, I want to be able to:
  - Catalog my data holdings
  - Aggregate data files
  - **Provide a ‘quick view’ of my data**
  - Easily add information (metadata) to my datasets
  - Fix ‘incorrect’ datasets*
  - Allow flexibility in the way users access my data

Give users what they need to do science!
Why TDS?

Data Providers

- As a data provider, I want to be able to:
  - Catalog my data holdings
  - Aggregate data files
  - Provide a 'quick view' of my data
  - Easily add information (metadata) to my datasets
  - Fix 'incorrect' datasets
  - Allow flexibility in the way users access my data

Give users what they need to do science!
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
- Advertise available datasets and services via catalogs
- Data access (subset) services:
  - OPeNDAP
  - OGC WMS and WCS
  - NCSS
- Data collection services
  - Aggregation
  - Point/station collection
- Metadata services
  - THREDDS Catalog XML
  - nclISO: ISO, UDDC, NcML

* Writing netCDF-4 requires the netCDF-C library
THREDDS Data Server

Servlet Container

THREDDS Server

NetCDF-Java Library

configCatalog.xml

catalog.xml

Remote Access Client

IDD Data

Datasets

thredds.ucar.edu
THREDDS Data Server
Getting Started

• TDS is written in 100% Java
• TDS uses the Java Servlet framework
  – Need to Install Tomcat or other servlet container
    • Tomcat used in many places (The Weather Channel, Netflix, LinkedIn, to name a few)
    • Note: many cloud services can use servlets (e.g. Amazon Web Services, CloudBees, Google App Engine, Windows Azure Compute, etc.)
THREDDS Data Server
Getting Started

• TDS is written in 100% Java

• TDS uses the Java Servlet framework
  – Need to Install Tomcat or other servlet container
    • Tomcat used in many places (The Weather Channel, Netflix, LinkedIn, to name a few)
    • Note: many cloud services can use servlets (e.g. Amazon Web Services, CloudBees, Google App Engine, Windows Azure Compute, etc.)

• First up: **Install and configure Tomcat**
  Docker and TDS