Creation of a case study archive using THREDDS

Leigh Orf
Marty Baxter
Central Michigan University

Motivation

- Case studies serve as a critical learning tool for students
- Wish to enable students to interrogate data (not just display pre-made imagery) and customize their output
- Want to serve classic cases indefinitely
- Wish to be able to easily add new cases as they occur (from ingested data on disk) and allow the repository to grow

Context

- Two MET faculty, with substantial UNIX experience and experience with Unidata software but no specific experience with THREDDS.
- We figured, “Serving data up with THREDDS seems to be the way to go for case study archives, so let's install it and see what happens”

Our approach

- Install THREDDS and learn about how it works (in that order, of course)
- Follow tutorial to get real-time ingestion working
- “Work backwards” to get COMET Case Study 1: Storm of the Century served via THREDDS
- Meet for an hour or two every Friday for a month or two

Installing and configuring THREDDS

- Took two meetings
- Install java, .war file
- Install tomcat – whoops next time don't do it as root (create tomcat user, reinstall as tomcat)
- Install three extra pqact.conf files – we ingest lots of duplicate data now (deal with this later)
- Ran into problems – needed Ethan to log in remotely, fix strange things, before we could access realtime data with IDV via THREDDS

Proof of Concept: Comet 001

- Now for the “work backwards” part: I want to serve up a piece of data (anything) via THREDDS.
- Need to muck with .xml files
- This is where I tell you “gee – a bunch of .xml templates for different file types would be really useful!”
- The dataType and dataFormat tags seem to be the key here – need to put the right magic here
- Can we create xml files describing: Model data, satellite data, surface data, upper air, and radar?

Proof of Concept: Comet 001

- Gridded data turns out to be easy, and with IDV reading .gem files, we're getting somewhere
- Radar: Had problems here. Could not get NCDC Level II data to work using templates which work for .ar2v files (what are these, anyway?)
- Satellite: Ended up having to install ADDE in order to get satellite to loop correctly
- Surface data: metar2nc worked
- Upper air: ua2nc did not work (well IDV couldn't grok the files created)

Our feedback

- We need to know what is and what is not do-able under THREDDS
- Having to get ADDE working to do satellite loops seems to defeat the purpose
- We need template .xml files for the standard case study data sources and/or clear documentation
- Virtual paths in .xml file are very confusing, can't tell where the data really is
- If we are to use THREDDS as a real-time data server, do we have to duplicate data so GEMPAK can still run happily?

Feedback concerning IDV

- In order for a case study to work, we must be able to display several different data types simultaneously
- Time matching becomes a problem. We would like IDV to behave like Gempak does for layering data that's not quite time synced
- We are still having performance and usability issues with IDV that make us nervous about having to depend on it as our only viewer for THREDDS

Final Comments

- Upon learning of RAMADAAAAAAAH! we wonder whether we shouldn't just scrap it and start over
- I am quite excited that Unidata is dedicating resources towards the next-gen case studies.
- It seems that some sort of convergence between these fledgling projects will lead to something that “just works”