



**Web-based Portal for
Discovery, Retrieval and Visualization
of Earth Science Datasets
in Grid Environment**

Zhenping (Jane) Liu



Outline

- ❖ **Challenges to share Earth Science Datasets**
- ❖ **Project Goals**
- ❖ **Proposed Solution**
- ❖ **Proposed System Architecture**
- ❖ **Demo System**

Challenges to share Earth Science Datasets

- ❖ **Difficulty caused by diverse data formats**
- ❖ **Difficulty to discover heterogeneous and distributed datasets**
- ❖ **Lack of data query and retrieval services**
- ❖ **Difficulty of data visualization and understanding**

Project goals

- ❖ **To provide a Web-based Portal for Discovery, Retrieval and Visualization of Earth Science datasets with extensibility, scalability, uniformity, transparency and heterogeneity in grid environments.**

Specific Project goals

- ❖ **For datasets sharing, implement**
 - ❧ **Dynamic Discovery**
 - ❧ **Heterogeneity Transparency**
 - ❧ **Location and Name Transparency**
 - ❧ **Distribution Transparency**
 - ❧ **Replication Transparency**
- ❖ **Remote and Interactive web-based visualization**
- ❖ **Thin Clients (Web browser)**

Proposed solution

- ❖ **Grid Technology**
- ❖ **Web Services Technology**
- ❖ **Java/J2EE**
- ❖ **Scientific Visualization Technology**
- ❖ **Four-tier Architecture**

A Layered View of Our System



Grid Technology

- ❖ **Controlled and coordinated sharing of geographically distributed, dynamic and heterogeneous resources.**
- ❖ **Grid Middleware**
 - ⌘ **Provide fundamental infrastructure for computing and data management.**
 - ⌘ **Permits application services to interface with the resources in a uniform way.**

Web Services Technology

- ❖ **Web Service is a platform and implementation independent software component that can be:**
 - ❧ **Described**
 - ❧ **Published**
 - ❧ **Discovered**
 - ❧ **Invoked**
 - ❧ **Composed with other services**

Benefits of Web Services

- ❖ **Reducing complexity by encapsulation**
- ❖ **Promoting interoperability**
 - ❧ Truly platform and language independent
- ❖ **Enabling interoperability of legacy applications**

Java/J2EE (JSP,Servlet,JavaBeans)

❖ **Web portal development**

Scientific Visualization

- ❖ **Represent huge amount of data graphically to help better understanding of the data**
- ❖ **Remote and Interactive Scientific Visualization**

Proposed System Architecture

❖ Four-tier Architecture

- ☞ Data Sources tier

- ☞ Grid Services tier

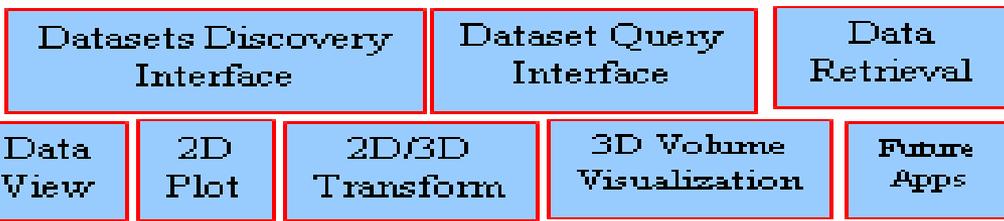
- ☞ Application Web Services tier

- ☞ Clients tier

Data Portal

Clients

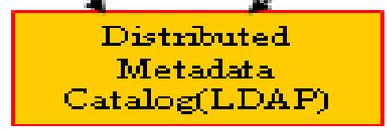
SOAP over Http, JavaScript, JavaApplet and etc



Application Web Services

XML/SOAP over Grid Security Infrastructure

Data Request Mgt Services



Grid Services



Grid Protocols and Grid Security Infrastructure



Data Sources

Grid Services Tier

❖ Issues addressed

- ❧ Resource Access and Management
- ❧ GSI Security Services
- ❧ High Performance Data Transport Services
- ❧ Metadata Catalog and management
- ❧ Replica Catalog and management

Grid Services Tier (Cont.)

❖ Distributed metadata catalog

- ❧ Stores physical and conceptual information of datasets
- ❧ Allows managing and accessing datasets intelligently and efficiently
- ❧ Plays a key role in the areas of managing, discovery and sharing of datasets.

Grid Services Tier (Cont.)

❖ Metadata management services

☞ Metadata query, search and discovery, extraction, conversion, aggregation, validation, registration, browsing, display, and metadata schema definition.

Grid Services Tier (Cont.)

- ❖ **Distributed Replica catalog**

- ❧ Provides mappings between logical names for files and the storage locations of one or more replicas of these files.

- ❖ **Replication management services**

- ❖ **Replication selection services**

Application Services Tier

- ❖ **Datasets discovery interface generation**
- ❖ **Data query interface generation**
- ❖ **Data Retrieval**
- ❖ **Data Viewer**
- ❖ **Scientific Visualization and Analysis**
 - ❧ **2D plot, 2D/3D Transform, 3D Volume Visualization.**
- ❖ **Future applications**

Clients Tier

❖ Web-based Data Portal

- ☞ All application services are delivered with web-browser
- ☞ Key advantages to thin clients

What've been done

- ❖ **A demo system: Web-based data management, retrieval, analysis and visualization system**
- ❖ **Implementing authentication and authorization web service module by using Globus Grid Security Infrastructure (GSI).**
- ❖ **Implementing access control web service module.**
- ❖ **Implementing data transfer service module by using GridFTP.**

Demo system – Features

- ❖ **Web-based portal**

- ❧ All application services are delivered with web browser.

- ❧ Thin clients

- ❖ **Several hundred of distributed earth science data sources are integrated into the system.**

- ❖ **Several common scientific data formats supported**

Demo system – Features (Cont.)

❖ Data management based on metadata mechanism

- ❧ Metadata to describe logical category of datasets
- ❧ Metadata to customize the query GUI for a dataset
- ❧ Metadata to describe logical directories (with content and semantic information) within a dataset
- ❧ Metadata to describe format and structure of a data file.
- ❧ Metadata to define available analysis methods for a dataset or a data file.

Demo system – Features (Cont.)

- ❖ **Dynamically generated dataset discovery web interface based on metadata**
 - 🌀 **Sample snapshots -- next two slides**

Select **Search:**

- [-] DODS Dataset List
 - [-] Data Providers
 - [-] Antarctic Cooperative Research Centre, Tasmanian Partnership for Advanced Computing (TPAC)
 - [-] Digital Library for Oceans and Climate
 - [-] **TPAC/CSIRO Climatologies**
 - [-] TPAC/CSIRO Satellite Altimetry
 - [-] NCEP - DOE Reanalysis 2
 - [-] WOCE Global Data Version 3.0
 - [-] Australian Antarctic Automatic Weather Station Dataset
 - [-] Carolinas Coastal Ocean Observing and Prediction System (Caro-COOPS)
 - [-] Storm Surge Data
 - [-] Center for Ocean Land Atmosphere Studies (COLA)
 - [-] COLA AGCM Model Data
 - [-] COLA AVN Model Data
 - [-] COLA Eta Model Data
 - [-] COLA MRF Model Data
 - [-] Columbia University/LDEO - International Research Institute (IRI/LDEO)
 - [-] ARCTIC
 - [-] Biosphere - A Global Change Laboratory

Selected Datasets

TPAC/CSIRO Climatologies
dir: <http://www.marine.csiro.au/dods/nph-dods/dods-data/climatology-netcdf/>

Demo system – Features (Cont.)

- ❖ **Dynamically generated data query web interface based on metadata**
 - 🌀 **Sample snapshots -- next two slides**

Directory listing for http://caro-coops.org/cgi-bin/nph-dods/storm_surge_data/

- http://caro-coops.org/cgi-bin/nph-dods/storm_surge_data/
 - bathy
 - buoys
 - current_discharge ...
 - discharge ...
 - noaa ...
 - wind ...

1 files:

- [usc_MET1_met_latest.nc](#)

time: time=[0:1:104]

Grid {

- source_time: time=[0:1:104]
- time: time=[0:1:104]

} source_time;

date_time_string: time=[0:1:104] ndate=[0:1:18]

Grid {

- latitude: time=[0:1:104]
- time: time=[0:1:104]

} latitude;

Grid {

- longitude: time=[0:1:104]
- time: time=[0:1:104]

} longitude;

Grid {

- wind_speed: time=[0:1:104]
- time: time=[0:1:104]

} wind_speed;

Grid {

- wind_from_direction: time=[0:1:104]
- time: time=[0:1:104]

Please select the conditions for character fields:

Basin Centroids (Unit:):

Please select the conditions for query fields:

Date (Format: "YYYY MM DD HH"): From To

Air temperature (Unit: K): From To

Surface Pressure (Unit: Pa): From To

Downward solar radiation (Unit: W/m^{**2}): From To

Downward long-wave radiation (Unit: W/m^{**2}): From To

Specific humidity (Unit: Kg/Kg): From To

U wind speed (positive means from west to east) (Unit: m/s): From To

V wind speed (positive means from south to north) (Unit: m/s): From To

Please customize the result format:

Date:

Air temperature:

Surface Pressure:

Downward solar radiation:

Demo system – Features (Cont.)

- ❖ **Efficient data retrieval based on metadata**
- ❖ **Web-based data browser**
 - 🌀 **Sample snapshots -- next two slides**

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites Media Print Copy Paste

Address <http://208.17.194.105:8080/HyDRO/dataDisaply.jsp> Go Links

Query Results

Year	Month	Day	Hour	Tair	Psfc	DSWR	DLWR	SPFH	U	V
1997	1	1	0	292.560	101391.	0.00000	372.180	0.0130000	12.3100	0.160000
1997	1	1	1	292.280	101451.	0.00000	370.620	0.0125000	10.4200	-2.23000
1997	1	1	2	291.440	101512.	0.00000	369.070	0.0120000	8.53000	-4.62000
1997	1	1	3	290.590	101572.	0.00000	367.510	0.0115000	6.64000	-7.00000
1997	1	1	4	289.750	101633.	0.00000	358.460	0.0110000	4.76000	-9.39000
1997	1	1	5	288.900	101693.	0.00000	349.420	0.0106000	2.87000	-11.7700
1997	1	1	6	288.790	101753.	0.00000	340.370	0.0101000	0.980000	-14.1600
1997	1	1	7	288.230	101867.	0.00000	331.320	0.00990000	0.580000	-13.8300
1997	1	1	8	287.660	101981.	0.00000	322.270	0.00960000	0.180000	-13.5100
1997	1	1	9	287.100	102094.	0.00000	313.220	0.00940000	-0.230000	-13.1800
1997	1	1	10	286.540	102208.	0.00000	311.550	0.00920000	-0.630000	-12.8500
1997	1	1	11	285.970	102322.	54.2700	309.870	0.00900000	-1.03000	-12.5300
1997	1	1	12	286.510	102435.	157.730	308.200	0.00870000	-1.43000	-12.2000
1997	1	1	13	285.960	102438.	250.590	306.530	0.00860000	-1.48000	-11.3300
1997	1	1	14	285.410	102441.	360.970	304.860	0.00840000	-1.53000	-10.4600
1997	1	1	15	284.860	102444.	420.450	303.180	0.00820000	-1.58000	-9.59000
1997	1	1	16	284.310	102446.	451.440	305.100	0.00800000	-1.63000	-8.71000
1997	1	1	17	284.600	102449.	443.090	307.020	0.00780000	-1.68000	-7.84000
1997	1	1	18	285.150	102452.	413.420	308.930	0.00760000	-1.73000	-6.97000
1997	1	1	19	284.850	102458.	355.720	310.850	0.00760000	-2.37000	-5.95000
1997	1	1	20	284.930	102464.	239.920	312.760	0.00770000	-3.02000	-4.93000
1997	1	1	21	285.230	102471.	152.080	314.680	0.00770000	-3.67000	-3.90000
1997	1	1	22	285.530	102477.	54.0100	314.680	0.00780000	-4.31000	-2.88000
1997	1	1	23	285.830	102483.	0.00000	314.680	0.00780000	-4.96000	-1.86000

Demo system – Features (Cont.)

- ❖ **Web-based remote and interactive 2-D/3-D data visualization toolkits**

- ℳ **2-D**

- ❖ **Plots, Colormaps, Contours**

- ℳ **3-D**

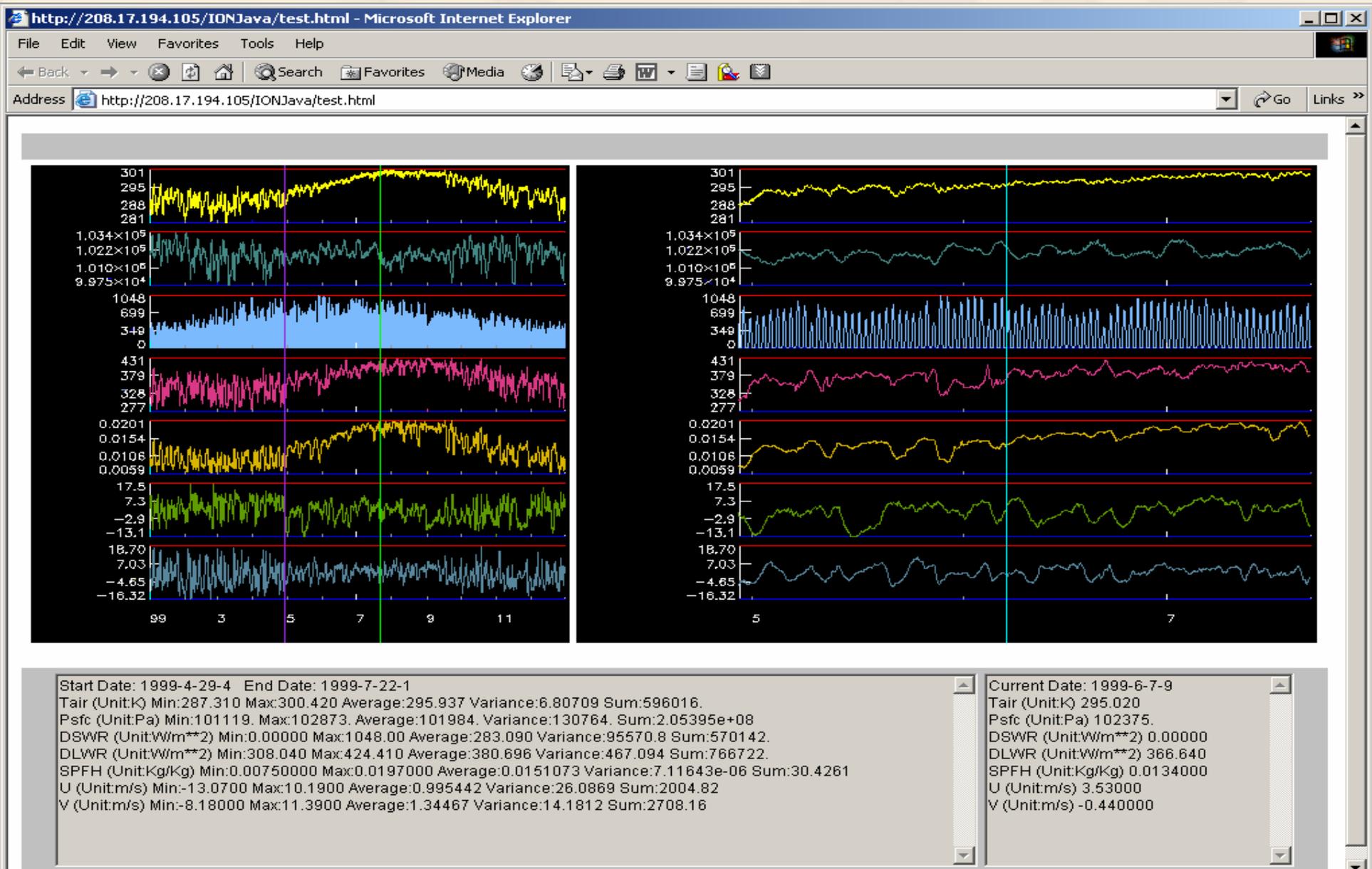
- ❖ **Surface**

- ℳ **Animation**

- ❖ **Plots, Colormap, Contours, 3-D Surface**

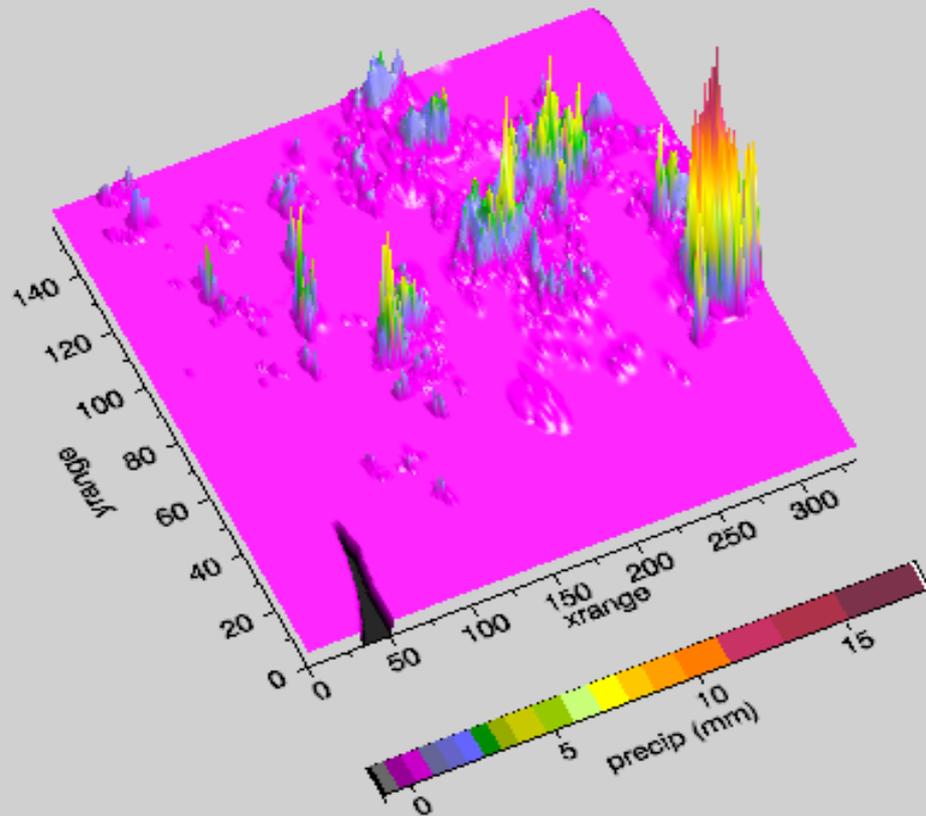
- ℳ **Sample snapshots (See next slides)**

Sample – 2-D Plots



Sample – Surface

precip (mm)
Date: 1995-05-01-00



View Options

3D

Color Bar

Surface

Shaded

Skirt

Flat

Gray

Contour Lines

Off

Color

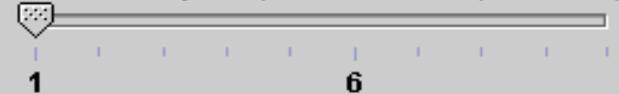
Off

Animation Action

Play

Stop

Animation Speed (Frames Interval, Unit :Sec)

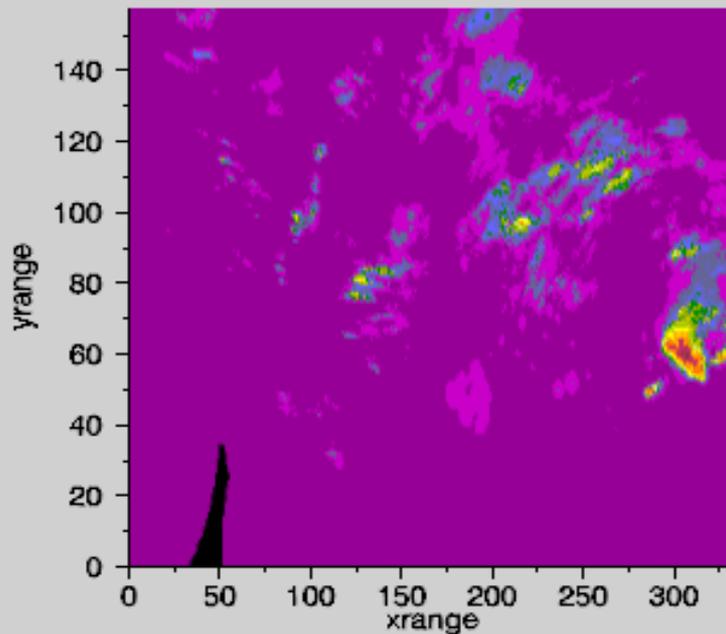


Animation Frame



Sample – 2-D Colormap

precip (mm)
Date: 1995-05-01-00



View Options

3D

Color Bar

Surface

Shaded

Skirt

Flat

Gray

Contour Lines

Off

Color

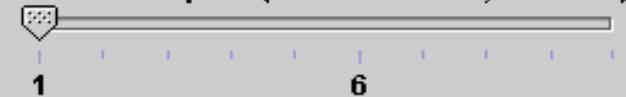
Off

Animation Action

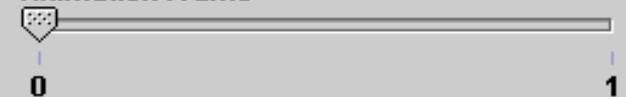
Play

Stop

Animation Speed (Frames Interval, Unit :Sec)

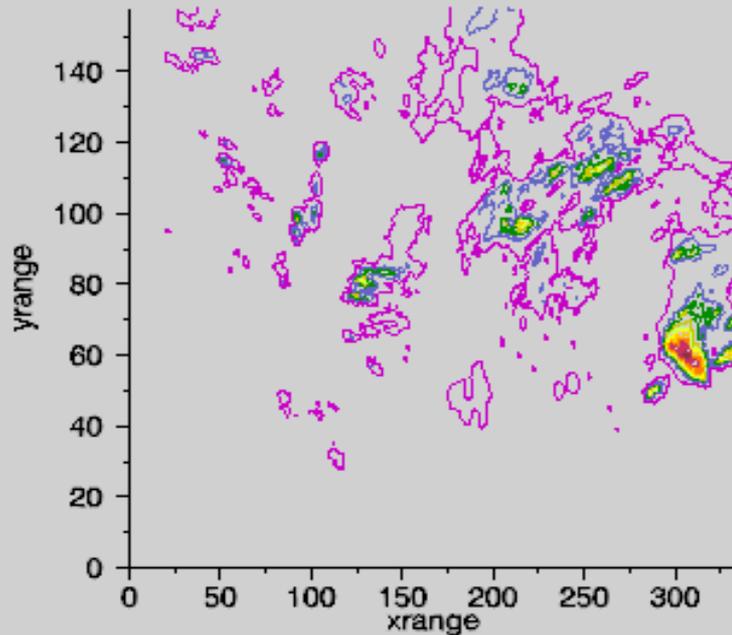


Animation Frame



Sample – 2-D Contour

precip (mm)
Date: 1995-05-01-00



View Options

3D

Color Bar

Surface

Off

Skirt

Flat

Gray

Contour Lines

Top

Color

Top

Animation Action

Play

Stop

Animation Speed (Frames Interval, Unit :Sec)



1

6

Animation Frame



0

1

GSI Authentication and Authorization web service

- ❖ **Primary security mechanism in the system.**
- ❖ **Data Portal retrieve a proxy certificate from a MyProxy server and act on users' behalf.**

Data Transfer web service

- ❖ **Allows a file to be transferred between two locations using one of several transport protocols: filesystem I/O, HTTP, FTP, HTTPS, or GridFTP.**
- ❖ **In the case of GridFTP, a credential is first retrieved from a MyProxy server and used to authenticate to the GridFTP server.**

Access control web service

- ❖ Gets a list of access privileges of the user after querying the access rights of the user from the database.

Project web site

❖ Project Introduction

🔗 <http://filebox.vt.edu/eng/ece/dmv/Grid/index.htm>



*Thank
You*

